

Establishment of Genome Editing Methods and Mechanistic Dissection of the Role of *fallen angel* during Sarcomere Maturation in *Drosophila*

**Dissertation der Fakultät für Biologie der
Ludwig-Maximilians-Universität München**



Xu Zhang

Max-Planck-Institut für Biochemie

Deutschland, 2017

This work has been conducted in the laboratory of Dr. Frank Schnorrer at the Max Planck Institute of Biochemistry (Munich, Germany) from September 2012 to December 2016.

First reviewer: Prof. Dr. Barbara Conradt

Second reviewer: Dr. Frank Schnorrer

Date of submission: 05.04.2017

Date of defence: 05.07.2017

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die von mir vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt ist.

München, den 05.07.2017

Xu Zhang
(Unterschrift)

Erklärung

Hiermit erkläre ich,

☒ dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist.

☒ dass ich mich anderweitig einer Doktorprüfung ohne Erfolg nicht unterzogen habe.

☐ dass ich mich mit Erfolg der Doktorprüfung im Hauptfach und in den Nebenfächern bei der Fakultät für der

(Hochschule/Universität)

unterzogen habe.

☐ dass ich ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen.

München, den 05.07.2017

Xu Zhang
(Unterschrift)

Contents

Contents	i
Abbreviations	iii
List of publications.....	vii
Declaration of contributions	viii
Summary.....	ix
Zusammenfassung.....	xi
1 Introduction	1
1.1 Skeletal muscles	2
1.1.1 Overview of muscle structure	2
1.1.2 Sliding-filament theory and cross-bridge model of muscle contraction.....	4
1.1.3 Regulation of muscle contraction	6
1.1.3.1 Calcium and Troponin/Tropomyosin complex	6
1.1.3.2 Myosin regulatory light chain	7
1.1.3.3 Stretch activation	7
1.2 <i>Drosophila</i> muscle system	8
1.2.1 Overview of <i>Drosophila</i> myogenesis	8
1.2.1.1 <i>Drosophila</i> embryonic body myogenesis	8
1.2.1.2 <i>Drosophila</i> adult myogenesis	9
1.2.1.3 <i>Drosophila</i> adult muscle types.....	9
1.2.2 Development of <i>Drosophila</i> indirect flight muscles	11
1.2.2.1 Myoblast specification	11
1.2.2.2 Founder cell specification	11
1.2.2.3 Myoblast fusion	12
1.2.2.4 Muscle tendon attachment	13
1.3 Myofibrillogenesis	13
1.3.1 The models of myofibrillogenesis	13
1.3.1.1 The two historical models	13
1.3.1.2 A tension driven model of myofibrillogenesis.....	14
1.3.2 Myofibril and sarcomere maturation in <i>Drosophila</i> IFM.....	16
1.4 Transcriptional regulation of IFM specification and differentiation	16
1.4.1 Mef2 - A general factor	17
1.4.2 Identity genes of IFM patterning and differentiation	17
1.4.2.1 Spalt major (Salm)	17
1.4.2.2 Extradenticle (Exd) and Homothorax (Hth)	18
1.4.3 The downstream targets of Salm	19
1.4.4 Splicing regulation during IFM development	20
1.5 IFMs as a model to study human muscle hypercontraction.....	20

1.5.1	Conservation of muscle biology from fly to human.....	20
1.5.2	Model for muscle hypercontraction.....	21
1.6	Genetic <i>Drosophila</i> resources.....	22
1.6.1	RNAi libraries, Fosmid libraries, and MiMiC collections	22
1.6.2	Precise genome editing methods: TALENs and CRISPR.....	23
1.7	Aims of the thesis.....	25
2	Discussion and Outlook	27
2.1	Establishment of precise genome editing methods in <i>Drosophila</i>	27
2.1.1	Applications of TALENs in <i>Drosophila</i>	27
2.1.2	CRISPR/Cas9	28
2.1.2.1	General applications of CRISPR/Cas9 in <i>Drosophila</i>	28
2.1.2.2	Combination of CRISPR/Cas with RMCE	28
2.2	<i>Fallen angel (fall)</i> , a novel BTB-ZF transcription factor, is essential for sarcomere growth and maturation.....	30
2.2.1	The <i>fall</i> loss-of-function phenotype	30
2.2.1.1	<i>fall</i> loss-of-function flies are viable but flightless	30
2.2.1.2	Sarcomere maturation defect and hypercontraction in <i>fall</i> mutants ...	30
2.2.2	Fall regulates the transcription of sarcomeric genes.....	31
2.2.2.1	Fall is a transcriptional regulator	31
2.2.2.2	Sarcomere maturation defect and hypercontraction might be caused by misregulation of Fall targets	32
2.2.3	Fall expression and localization pattern	33
2.2.3.1	Body muscle specific Fall expression and function.....	33
2.2.3.2	Fall localises in nuclear body pattern in IFM	33
2.2.3.3	Structural function analysis of Fall	34
2.2.4	BTB-zinc finger proteins and conservation.....	35
2.2.4.1	BTB-ZFs are conserved across organisms.....	35
2.2.4.2	Nuclear body formation: a common feature of BTB-ZF proteins?	36
2.2.4.3	Fall: a homolog of ZBTB42 in fly?	37
2.3	Outlook.....	37
	References:.....	39
	Appendices.....	49
	Publication I	51
	Publication II.....	56
	Manuscript I	67
	Acknowledgement	112
	Curriculum Vitae	113

Abbreviations

ADP	Adenosine diphosphate
AME	Adult muscle enhancer
AMP	Adult muscle progenitor
APF	After puparium formation
Aret	Arrest
Arp2/3	Actin related protein 2/3
ATP	Adenosine triphosphate
attP	Attachment site on phage
attB	Attachment site on bacteria
attR	Attachment site on right
BEAF-32	Boundary element-associated factor of 32kD
BTB	Broad-complex, Tramtrack and Bric-a-brac
CDS	Coding DNA sequence
CELF	CUG-binding protein and ETR-3-like factor
ChIP-seq	Chromatin immunoprecipitation sequencing
CP190	Centrosome-associated zinc finger protein CP190
CRISPR	Clustered regularly interspaced short palindromic repeats
Ct	Cut
CTCF	CCCTC-binding factor
DamID	DNA adenine methyltransferase identification
DESeq	Differential expression analysis for sequence count data
DEXSeq	Differential usage of exons from RNA-seq data
DFM	Direct flight muscle
DLM	Dorsal longitudinal flight muscle
DSB	Double strand break
Duf	Dumbfounded
DVM	Dorsal ventral flight muscle

ELC	Essential light chain
Exd	Extradenticle
F-actin	Filamentous actin
Fall	Fallen angel
FC	Founder cell
FCM	Fusion competent myoblast
Fhos	Formin homology 2 domain containing
Fln	Flightin
FRT	Flippase recognition target
G-actin	Globular actin
GFP	Green fluorescent protein
Hbs	Hibris
HDR	Homology directed repair
HRMA	High resolution melt analysis
Hth	Homothorax
L'sc	Lethal of scute
LCCS	Lethal congenital contractural syndrome
LOM	Larval oblique muscles
Mef2	Myocyte enhancer factor-2
MHC	Myosin heavy chain
MiMIC	Minos mediated integration cassette
MLCK	Myosin light chain kinase
MLCP	Myosin light chain phosphatase
Mod(mdg4)67.2	Modifier of mdg4 67.2kD
MTJ	Muscle tendon junction
MYBPC1	Slow skeletal muscle myosin binding protein C
MYH3	Embryonic myosin heavy chain
MYH8	Fetal myosin heavy chain
NcoR	Nuclear receptor corepressor
NHEJ	Non-homologous end joining

NIG-FLY	Fly stocks of national institute of genetics
P bodies	Processing bodies
PAM	Protospacer adjacent motif
PCR	Polymerase chain reaction
PloyN	Polyasparagine
PML	Promyelocytic leukemia
PolyQ	Polyglutamine
RLC	Regulatory light chain
RMCE	Recombinase-mediated cassette exchange
RNAi	RNA interference
Rst	Roughest
Salm	Spalt major
Salr	Spalt related
Sals	Sarcomere length short
SC-35	SR family splicing factor SC35
sgRNA	Single strand guide RNA
Sls	Sallimus
SMRT	Silencing mediator for retinoid and thyroid hormone receptor
Sns	Sticks and stones
Sr	Stripe
Strn-mlck	Stretchin-Myosin light chain kinase
Stv	Starvin
Su(Hw)	Suppressor of hairy wings
TALENs	Transcription Activator-like Effector Nucleases
TNNI2	Fast troponin I
TNNT3	Fast troponin T
TPM2	Beta-tropomyosin
Tpn C	Troponin C
TRiP	Transgenic RNAi project
UAS	Upstream activation sequence

Vg	Vestigial
WASp	Wiskott-Aldrich syndrome protein
ZASP52	Z band alternatively spliced PDZ-motif protein 52
ZASP66	Z band alternatively spliced PDZ-motif protein 66
ZBTB42	Zinc finger and BTB containing protein 42
ZF	Zinc finger
α PS	Position-specific integrin subunit alpha
β PS	Position-specific integrin subunit beta

List of publications

Publications:

1. Zhang X, Ferreira IRS and Schnorrer F (2014) A simple TALEN-based protocol for efficient genome editing in *Drosophila*. *Methods* 69, 32–37.
2. Zhang X, Koolhaas WH and Schnorrer F (2014) A Versatile Two-Step CRISPR- and RMCE-Based Strategy for Efficient Genome Engineering in *Drosophila*. *G3: Genes| Genomes| Genetics* (Bethesda) 4, 2409–2418.

Manuscript:

1. Zhang X, Spletter M, Habermann B and Schnorrer F (2017) The BTB-Zinc finger *fallen angel* is essential for correct sarcomere maturation to prevent muscle hypercontraction. Manuscript ready for submission.

Declaration of contributions

Publication 1:

X.Z. performed all experiments and IRS.F. contributed to fly sample collection. X.Z. and F.S. analysed data and wrote the manuscript. F.S. conceived the study.

Publication 2:

X.Z. established the CRISPR-RMCE based strategy and characterised *salm* knockout/knockin alleles. W.K. generated *bt* allele. X.Z. and F.S. analysed data and wrote the paper. F.S. supervised the study.

Manuscript 1:

X.Z. performed all the wet experiments. B.H. analysed the RNA-seq data. M.S. and X.Z. integrated the RNA-Seq with proteomics data and made figure 6. X.Z. and F.S. generated all other figures and wrote the manuscript. F.S. conceived and supervised the project.

I hereby confirm the above statements being correct.

Xu Zhang

Dr. Frank Schnorrer

Summary

All animals, from simple flies to humans, rely on a functional muscle system to produce and coordinate their sophisticated movements. Our muscle system has evolved to perform specific behaviors such as feeding, mating, and laughing among many others. Interestingly, apart from the general myogenic regulatory factors, many specialized transcription factors are recruited to cooperate with the core transcription factors to instruct the different developmental processes building different muscle types. However, how such a cooperation works, in particular how it can instruct myofibrillogenesis and sarcomerogenesis in a muscle specific way is largely unknown. To tackle this problem, I used a well-established model, the *Drosophila melanogaster* indirect flight muscles (IFMs), to explore the function of new players that instruct muscle development, which may discover novel principles potentially conserved to human muscle biology.

In the first part of the thesis, I have successfully established TALEN induced mutagenesis in *Drosophila*. With TALEN, I generated mutants for several candidate genes, including *CG11617*, *salr*, and *CG32121* (*fallen angel*, *fall*). We found that a loss of function allele of *fall* causes a flightless phenotype, while the flies are viable. Surprisingly, mutations in *CG11617*, a previously uncharacterized putative transcription factor, did not result in any obvious muscle phenotype in *Drosophila*.

In the second part I took genome editing a step further and developed an efficient two-step genome engineering strategy by combining CRISPR with recombinase-mediated cassette exchange (RMCE) and successfully adapted it to mutate and tag *salr* as well as *fall*. We standardized this protocol and applied it to manipulate many other favorite candidates genes in our lab at high efficiencies.

In the third part of my thesis, I could show that Fall, a novel BTB-zinc finger protein, is essential for correct IFM development. Null alleles of *fall* lead to hypercontracted muscle fibers resulting in rupturing of the muscle-tendon system and muscle atrophy. We found that Fall protein is capable to form nuclear bodies, which are distinct from insulator bodies. Structure-function analysis of Fall showed that both BTB domain and zinc finger domains are essential for Fall to regulate downstream targets, while only the

BTB domain is important for Fall to enter nucleus and form nuclear bodies. Most importantly, we found that Fall regulates the expression of many sarcomeric components specifically required for sarcomere maturation in order to enable normal sarcomere growth during flight muscle development. Our results suggest that myofibrillogenesis initiation and sarcomere maturation are regulated by a distinct set of transcriptional regulators ensuring proper formation of the final pseudo-crystalline sarcomeric organization of the flight muscles.

Zusammenfassung

Alle Tiere, ob kleine Fliegen oder komplexe Menschen, benötigen ein funktionelles Muskelsystem für ihre ausgeklügelten Bewegungen. Unser Muskelapparat hat sich entwickelt, um ein Vielzahl von Verhaltensweisen, wie Fressen, Lachen oder Sex unterstützen zu können. Für die korrekte Entwicklung werden interessanterweise neben den generellen Muskelschalergen eine Reihe von spezialisierten Transkriptionsfaktoren rekrutiert, um mit den Kern Transkriptionsfaktoren zu kooperieren. Auf diese Weise werden die verschiedenen Entwicklungsschritte, die zur Entwicklung der unterschiedlichen Muskelarten führen, gesteuert. Wie genau die Kooperation der verschiedenen Faktoren funktioniert, um die Myofibrillen und Sarkomerbildung zu steuern, ist unbekannt. Ich habe das etablierte Model der indirekten Flugmuskulatur von *Drosophila melanogaster* ausgewählt, um dieses Problem zu anzugehen und die Funktion von neuen Regulatoren der Muskelentwicklung zu studieren. Dies könnte zur Entdeckung neuer Mechanismen führen, die auch für die Muskelbiologie des Menschen wichtig sein könnten.

Im ersten Teil meiner Arbeit habe ich die TALEN induzierte Mutagenese in *Drosophila* etabliert. Mittels TALENs habe ich Mutanten in mehreren Kandidatengen, inklusive *CG11617*, *salr* und *CG32121* (*fallen angel*, *fall*) hergestellt. Wir haben herausgefunden, dass ein Funktionsverlust von *fall* zu lebenden Fliegen führt, die allerdings völlig fluglos sind. Erstaunlicherweise führen Mutationen in *CG11617*, einem vorher nicht studierten wahrscheinlichen Transkriptionsfaktor, zu keinem offensichtlichen Phänotyp.

Im zweiten Teil meiner Arbeit habe ich das Genome Editing einen Schritt weiter gebracht und habe eine effiziente Zweischnitt-Strategie für die gezielte Genveränderung entwickelt. Für diese Strategie habe ich CRISPR mit recombinase mediated cassette exchange (RMCE) kombiniert und erfolgreich angewandt, um damit Mutationen in *salr* und *fall* herzustellen und einen "Tag" in diese Gene einzubauen. Wir haben diese Protokoll weiter standardisiert und für die gezielte Veränderung einer ganzen Reihe von anderen Genen im Labor mit hoher Ausbeute angewandt.

Im dritten Teil meiner Arbeit konnte ich zeigen, dass *Fall*, ein neuer BTB-Zinkfinger Transkriptionsfaktor essentiell für die normale Flugmuskel Entwicklung ist. Genetische

Nullallele in *fall* führen zu Muskel Hyperkontraktionen und einem Zerreißen des Muskel-Sehen Systems, mit Muskelatrophy als Folge. Wir haben entdeckt, dass das Fall Protein im Zellkern zu Strukturen lokalisiert, die verschieden von den "Insulatorkörpern" sind. Mittels Struktur-Funktionsanalyse von Fall konnten wir zeigen, dass sowohl die BTB, als auch die Zinkfinger Domäne von Fall benötigt werden, um nachgeschaltete Zielgene zu regulieren. Dabei ist nur die BTB Domäne für den Kernimport und die Lokalisierung zu den Kernstrukturen notwendig. Die wichtigste Erkenntnis ist, dass Fall die Expression von vielen Sarkomer Bausteinen reguliert, die spezifisch für die Reifung der Sarkomere benötigt werden. Zusammen erlaubt dies ein normales Sarkomerwachstum während der Flugmuskelentwicklung. Unsere Ergebnisse legen nahe, dass der Beginn der Myofibrillen Bildung und das Reifen der Sarkomere von unterschiedlichen Sätzen an transkriptionellen Regulatoren gesteuert wird. So wird sicher gestellt, dass die pseudo-kristalline Organisation der Sarkomere in den Flugmuskeln richtig ausgebildet wird.

1 Introduction

A highly coordinated muscle system is indispensable for the daily life of animals, including simple insects and complex humans. During development, various muscles with different properties, shapes, and functions are built for their special needs at defined positions of our body. Muscle specialization is exemplified by comparing the muscles from a 100 meter runner to a marathon runner. The former ones are specialized to perform high mechanical work at the expense of endurance, while the latter are designated for long endurance workload. Imagining that the eyelid blinks once every minute, while the heart beats every second, one can easily appreciate the finely designed and coordinated muscle systems in the human body. Interestingly, the basic code, namely myogenic regulatory factors such as MyoD, Myf5, myogenin and MRF4 are shared by many different muscle types in different organs (Buckingham, 2001; 2006). How these core transcription factors are combined with muscle-type specific regulators to instruct functional muscle specialization is less clear.

The fruit fly, *Drosophila melanogaster*, has been used as a model for developmental biology and genetics over a century. Tremendous discoveries made in flies established the biological similarity between *Drosophila melanogaster* and mammals (Wangler et al., 2015). And not surprisingly, flies and mammals share many features in muscle development (Taylor, 2006). First, more than 60% of the 13,000 proteins that are encoded by the fly genome showed sequence similarity with human proteins, which indicates the fundamental similarity between *Drosophila* and vertebrate in general (Rubin, 2001). Second, in vertebrates, three categories of muscles, including skeletal, heart, and smooth muscles have been classified according to their structural and functional properties. Although the physiological characters differ, the somatic, the heart, and the visceral musculature of *Drosophila* are the anatomical and functional counterparts of these three classes. Finally, the mechanisms underlying muscle development show that the main processes of muscle formation including determination of myogenic identity, specification of muscle precursors, myoblast fusion, myofibrillogenesis, and sarcomerogenesis require the coordinated actions of evolutionarily conserved genes in both *Drosophila* and humans. The large variety of

genetic tools and its manipulation make the *Drosophila* muscular system, especially the adult muscles, an excellent model to dissect mechanisms of muscle morphogenesis.

1.1 Skeletal muscles

1.1.1 Overview of muscle structure

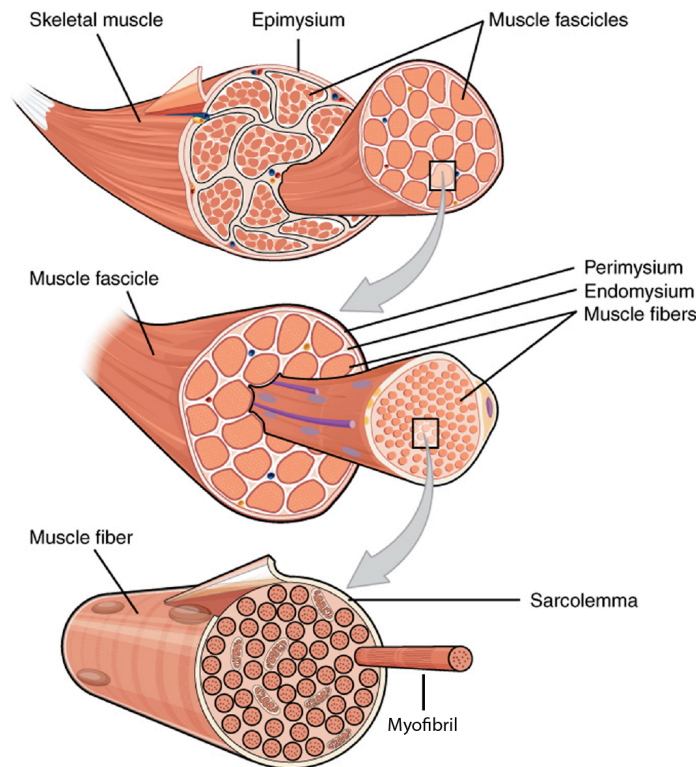


Figure 1: Scheme of skeletal muscle morphology (modified from Anatomy and Physiology, an OpenStax resource: <http://philschatz.com/anatomy-book/contents/m46476.html>)

In vertebrates, muscles are categorized into three types: skeletal, heart and smooth muscles. They share prominent common properties such as tissue contraction apart from diverse physiological functions. Here I take skeletal muscles as an example to illustrate the basic morphology of muscle tissues (Figure 1). Skeletal muscles generally form a cylindrical, spindle-like tissue, which is linked to the skeleton by connective tissue. Muscles are excitable and can contract upon neuronal stimulation. Each muscle is encapsulated in a sheath of connective tissue called epimysium that maintains muscle integrity during contraction by separating it from surrounding tissues. The inside space of the epimysium is filled with dozens of muscle fascicles, which are composed of organized muscle fiber bundles and are incased by a perimysium. Each muscle fiber is a

Introduction

muscle cell, which is a multinucleated syncytial cell. Inside of the muscle fibers are well-aligned hundreds to thousands myofibrils.

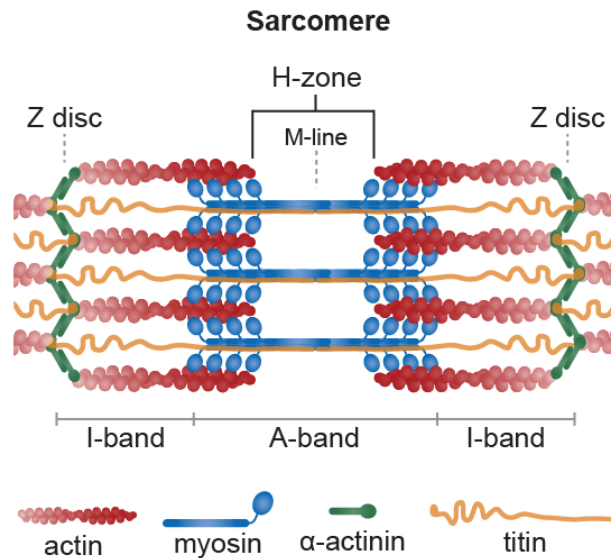


Figure 2: Sarcomere structure (modified from (Lemke and Schnorrer, 2016)).

M-line lies in the middle of a sarcomere. α -actinin localizes to Z disc and anchors actin filaments. The main components of sarcomere, actin based thin filaments and myosin based thick filaments, are interdigitated. The I-band spans the region that without thick filaments, while the A-band spans the region occupied by thick filaments. The H-zone is the thin filaments free region. Titin connects thin and thick filaments.

The most striking appearance of myofibrils is cross-striation, a pattern characterized by the alternating light (I-band, isotropic in polarized light) and dark bands (A-band, anisotropic in polarized light) observed under light microscope. This feature is directly caused by the highly ordered organization of the myofibrils, the mini-machines called sarcomeres, and their repetitive alignment along the myofibrils in a longitudinal manner (Clark et al., 2002). The fundamental contractile unit of each muscle fiber is the sarcomere. It is made of numerous sophisticated macromolecular protein complexes. Among them are the actin based thin filaments and myosin based thick filaments, as well as titin based elastic filaments (Figure 2). Each sarcomere is centered around the M-line and bordered by two Z discs, to which the thin filaments are cross-linked. The thin filaments originate from both Z discs, span the I-band, and point towards the M-line. They overlap with thick filament in the A-band but do not reach the M-line. The thin filaments free space is defined as H zone. The bipolar thick filaments are centered at the M-line and define the A band. They point towards the actin filaments, with which they overlap (Ehler and Gautel, 2008).

These actin and myosin based complexes form an almost crystalline order. The relative position and arrangement between the two kinds of filaments were resolved by electron microscopic studies. In longitudinal sections of muscle fibers, the thin and thick filaments align in parallel. Interestingly, in cross sections, a hexagonal array of thin filaments surrounding each thick filament was observed. These structural arrangements resulted in the proposition of the sliding filament theory of muscle contraction (Huxley, 2004).

1.1.2 Sliding-filament theory and cross-bridge model of muscle contraction

Skeletal muscles function mostly by producing mechanical force resulting in muscle contraction. A substantial body of studies has been performed to dissect the sarcomeric architecture and the mechanisms underlying muscle contraction. Decades ago, the sliding filament theory has been proposed. Earlier studies revealed that upon sarcomere contraction or stretching, the A-band length keeps constant and the edges of the H-zone move towards the Z disc. This observation indicates that during muscle contraction thin and thick filaments slide on each other (Hanson and Huxley, 1953; Huxley and Niedergerke, 1954; Squire, 2016).

Later investigations of the molecular composition of thin filaments and thick filaments largely expanded our understanding of the contraction mechanisms. The thin filament is a double twisted α -helix of polymerized actin molecules. These actin filaments are associated with troponin complexes, which are intertwined with tropomyosin that covers the myosin binding sites. Muscle myosin is a hexamer composed of two heavy chains (MHC), two essential light chains (ELC) and two regulatory light chains (RLC) and can be divided into two functional domains, a globular head and a rod. Each globular head is composed of the N terminus of the two MHC, two ELC and two RLC. It can form a cross-bridge to an actin filament. It possesses actin-activated ATPase activity and is capable to bind to actin. The myosin rod consists of the C terminal halves of the two heavy chains. The C terminal end of the rod contains the coiled-coil domains and is involved in myosin bipolar filament assembly. Between the motor head and the rod is the so-called lever-arm (Clark et al., 2002).

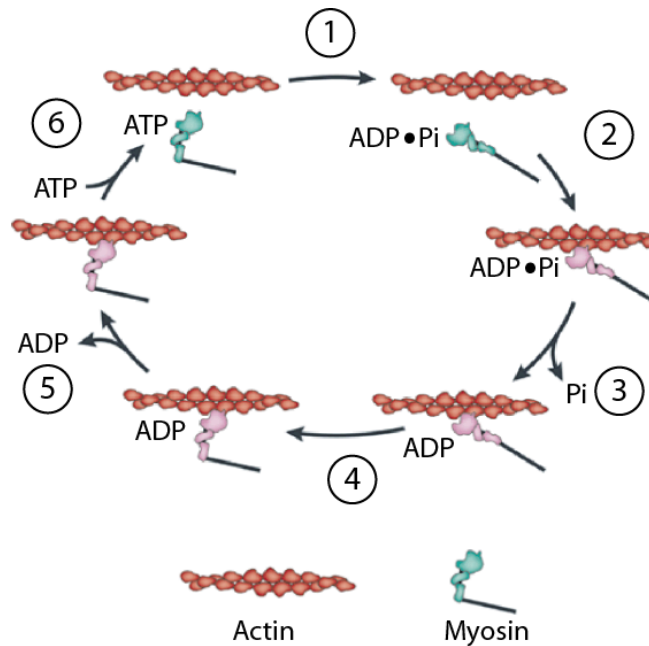


Figure 3: Myosin cross-bridge model, modified from J. A Spudich (Spudich, 2001). See details in the text.

In most scenarios, muscle contraction is initiated by neuronal stimulation. Upon neuronal stimuli, calcium channels open and release calcium from the sarcoplasmic reticulum into the sarcoplasm. Calcium binds to troponin, as a consequence tropomyosin moves and allows the myosin heads to attach to actin. Thus, the cross-bridge cycle is initiated and force is produced by ATP hydrolysis.

The swinging cross-bridge model could be dissected into six steps as illustrated in Figure 3. According to this model, the myosin head (the ATPase) spends most of its time detached from actin (step 1, 2). Once it is loosely attached to actin, the phosphate is released from the bound ADP•Pi and triggers the conformational change of the lever arm, which produces the power stroke (step 3, 4). This will move the actin filament towards the M line and increase the affinity between actin and myosin. This event will facilitate ADP release from the myosin head and enable the binding of the next ATP (step 5, 6). The binding of ATP to the myosin head domain will release the cross-bridge from the actin filament. Consequently, the ATPase of the myosin head is activated and induces ATP hydrolysis, which will cause a swing of the lever arm and prepare the next cycle. As each cycle only moves the actin filament 10 nanometers, the force generation for muscle contraction is a result of many of actinomyosin cycles (Geeves and Holmes, 1999; Gordon et al., 2000; Spudich, 2001).

1.1.3 Regulation of muscle contraction

In order to coordinate the complex movements of our daily life, muscle contractions have to be tightly regulated and finely tuned. The adjustment of muscle contraction is mostly reflected by the regulation of cross-bridge cycling. These regulations could be implemented in different ways, including the regulation of troponin complexes in skeletal and heart muscles, the phosphorylation of regulatory light chain by MLCK in smooth and heart muscle, and stretch activation in insect flight muscles (Gomes et al., 2002; Gordon et al., 2000; Linari et al., 2004; Ohtsuki and Morimoto, 2008; Sanders, 2008; Webb and Webb, 2003).

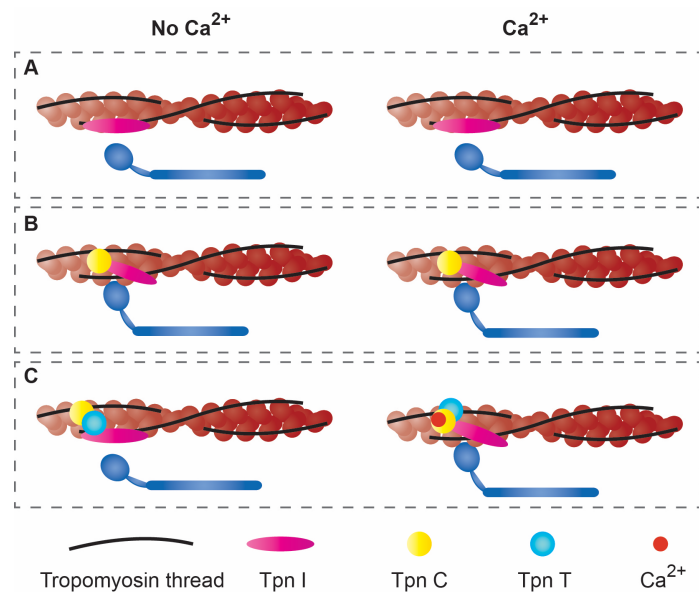


Figure 4: Regulation of muscle contraction by troponin and calcium, modified from (De Oliveira Vilaca, 2013).

A: Tpn I blocks the myosin binding site on actin filaments and this is not dependent on Ca^{2+} . B: Tpn C binds to Tpn I and release its blockage effect. C: Tpn T competes with Tpn I to bind to the Tpn I-binding site on Tpn C in the absence of Ca^{2+} . In the presence of Ca^{2+} , Tpn C undergoes a conformational switch and binds to Tpn I more preferably.

1.1.3.1 Calcium and Troponin/Tropomyosin complex

The troponin complex consists of troponin C, troponin I, and troponin T. Troponin C binds Ca^{2+} , troponin I is an inhibitory component while troponin T links the troponin complex and tropomyosin. Troponin I binds to the actin/tropomyosin threads and prevents the interaction of actin with myosin in the presence of tropomyosin (Figure 4A). Troponin C has a high affinity binding site for troponin I and can neutralize the suppressive effect of troponin I in the absence of troponin T by competing the binding site with the actin/tropomyosin threads. This desuppression of actinomyosin interaction

is independent of Ca^{2+} (Figure 4B). However, the existence of troponin T could prevent cross-bridge activation by masking the binding site of troponin I on troponin C (Figure 4C). Thus, Ca^{2+} plays an essential role in this case. Upon the binding of Ca^{2+} , troponin C undergoes a conformational switch and binds to troponin I in the presence of troponin T, thereby activating the cross-bridge cycle (Craig and Lehman, 2001; Gomes et al., 2002; Gordon et al., 2000; Ohtsuki and Morimoto, 2008).

1.1.3.2 Myosin regulatory light chain

Another regulatory mechanism involves the phosphorylation of myosin regulatory light chain, which can be implemented in two ways. In smooth muscle cells, increased Ca^{2+} levels activate myosin light chain kinase (MLCK) by the Ca^{2+} -calmodulin complex, thereby myosin regulatory light chain (RLC) could be phosphorylated by activated MLCK. Phosphorylated RLC activates the ATPase of myosin and thus increases muscle contraction (Takashima, 2009). It has been reported that phosphorylation of regulatory light chain also modulate skeletal, heart, and asynchronous insect flight muscle contraction (Farman et al., 2009; Kampourakis et al., 2016; Szczesna et al., 2002). Myosin light chain phosphatase (MLCP) can dephosphorylate the regulatory light chain, and thus block cross-bridge cycling. Thus, Rho-kinase activation triggers muscle contraction by inhibiting MLCP activity. This represents the second layer of regulation, which is independent of Ca^{2+} (Webb and Webb, 2003).

1.1.3.3 Stretch activation

Other types of muscles such as heart and asynchronous insect flight muscles can be activated by mechanical stretch, which represents another layer of regulation (Campbell and Chandra, 2006; Moore, 2006). This mechanism is particularly profound in asynchronous muscles such as insect flight muscle (Pringle, 1949). Different from synchronous muscles that contract at a 1:1 ratio with the neuronal spikes, asynchronous muscle contraction is not correlated with neuronal spiking (Josephson et al., 2000). Furthermore, asynchronous muscle contraction is not regulated by the oscillation of sarcoplasmic Ca^{2+} levels.

The molecular and regulatory mechanisms of stretch-activation are less well understood. Two hypotheses including the helix match-mismatch hypothesis and the strain sensor hypothesis have been proposed to explain stretch-activation (discussed by Hooper (Hooper et al., 2008)). It was believed that both of them might contribute to the stretch-

activation of asynchronous muscles. Recently, it has been shown that troponin bridges are involved in the regulation of stretch-activation in insect muscles. In the relaxed muscles, the myosin binding site on actin is blocked by tropomyosin. Upon stretch, troponin pulls on the troponin-tropomyosin complex and thus exposes the binding site on actin to the myosin head, resulting in muscle activation (Perz-Edwards et al., 2011). A second study showed that TnH (the larger IFM Tpn I isoform) is pulled off actin when the sarcomere is stretched. Tropomyosin is then released from the blocking position (Kržič et al., 2010). It is worth to note that in both cases, a permissive Ca^{2+} level is necessary for cross-bridges to bind to actin and to generate force.

1.2 *Drosophila* muscle system

1.2.1 Overview of *Drosophila* myogenesis

The *Drosophila* muscle system contains multiple types of muscles with different shapes, positions, sizes, and functions. Precise cooperation of these muscles allows flies to walk, fly, mate, and feed properly. *Drosophila* development shows two phases of myogenesis, one in the embryo that builds the larval musculature important for larval locomotion and feeding, and a second one during pupal development that forms all adult muscles, including the flight muscles (Roy and VijayRaghavan, 1999).

1.2.1.1 *Drosophila* embryonic body myogenesis

As in vertebrates, the *Drosophila* muscle system originates from the mesoderm layer after gastrulation in the early embryo. Initially, a combination of intrinsic transcriptional regulators and extrinsic signals are important to pattern the mesoderm into distinct domains in each embryo segment. A combination of signaling pathways establishes clusters of Lethal of scute (L'sc) expressing cells, called local equivalent groups, in the high Twist expression domain. Within each cluster of cells, one muscle progenitor is segregated by sustained expression of L'sc. Subsequently, these progenitors undergo one asymmetrically division to generate either two muscle founder cells (FCs) or one FC and one adult muscle progenitor cell (AMP) (Dobi et al., 2015; Ruiz Gomez and Bate, 1997; Schnorrer and Dickson, 2004). Other cells of the cluster will become fusion-competent myoblasts (FCM). These two different types, FCs and FCMs, are the source for embryonic body muscle myogenesis. In contrast, the AMPs that maintain high Twist expression remain undifferentiated until early pupal stage.

After the specification of FCs, FCMs migrate to FCs and establish contact with them by actin-based filopodia. These contacts will initiate fusion of FCMs to FCs (Schnorrer and Dickson, 2004). Consequently, the fusion of myoblast increases the muscle mass and the nuclei number of the myotubes, which will generate the final single-fiber syncytium (the so-called larval muscles) (Bate, 1990). Together, this equips each abdominal hemisegment of the embryo with 30 body muscles with distinct sizes, positions, morphology, tendon attachment, and innervation patterns.

1.2.1.2 *Drosophila* adult myogenesis

The second round of myogenesis takes place at the onset of metamorphosis. Most larval muscles start to be histolysed and the adult muscle are built *de novo*. The primary source of the adult musculature are stem cell-like AMPs that are born in the embryo (Figeac et al., 2010). During larval development AMPs migrate to the precursors of the adult tissues, such as wing and leg imaginal discs, with which they tightly associate and proliferate (Bate et al., 1991). In the pupae, these AMPs fuse with each other to form the adult muscle fibers *de novo* such as leg, abdominal, or dorso-ventral flight muscles (DVMs). In the exceptional case of the dorsal-longitudinal flight muscles (DLMs), AMPs fuse to remaining larval template muscles, which have escaped histolysis (Dutta et al., 2004). Similar to the larval muscles, after myoblast fusion has started, the myotube will contact the tendons to initiate attachments. Once attached, the initial myofibrils are being built and they grow to their final size (Weitkunat et al., 2014).

1.2.1.3 *Drosophila* adult muscle types

Drosophila adult body muscles constitute of two major types - fibrillar indirect flight muscles (IFMs) powering flight and tubular body muscles required for movement of most other body parts, such as head, legs or abdomen (Figure 5). The *Drosophila* tubular muscles are very similar to vertebrate body muscles as they are cross-striated with laterally aligned sarcomeres and contract synchronously, meaning every neuronal action potential will cause a fiber twitch (Figure 5).

Introduction

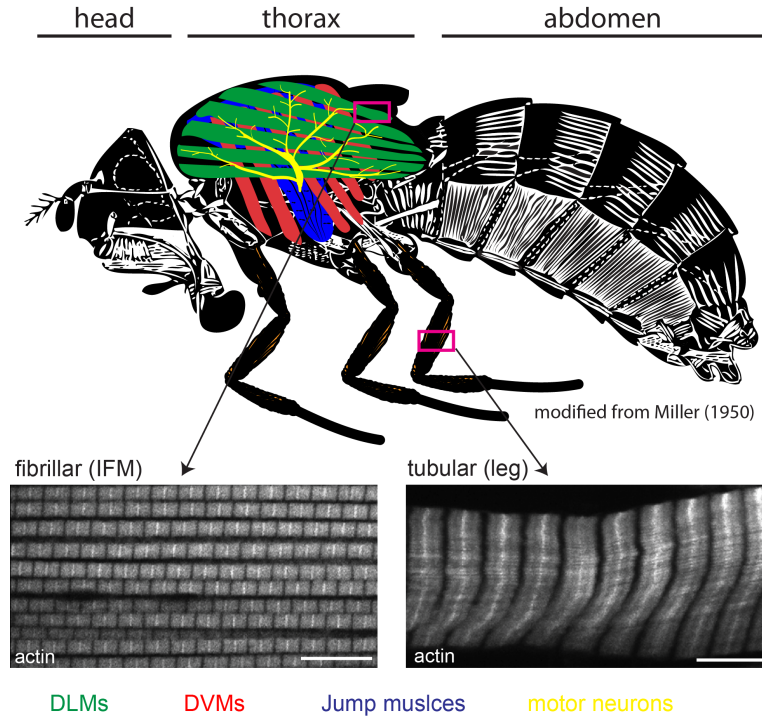


Figure 5: *Drosophila* adult muscle morphology, the fly scheme is modified from Manuela Weitkunat; The fibrillar and tubular muscles image are adapted from (Schönbauer et al., 2011).

The large fibrillar DLM (green fibers), DVM (red fibers), and tubular jump muscles (blue) are present in the thorax. Yellow branches show the motor neurons immigrating the flight muscles. Myofibrils in tubular leg muscles are well aligned with each other and show a cross-striation pattern, while in fibrillar IFMs are not aligned.

IFMs, the largest muscles in the thorax, are located in mesothorax and consist of two opposing muscle units, the DLMs and DVMs (Figure 5). Both units are not directly inserted at the wings but connected to the thorax, thus called indirect flight muscles. In sharp contrast to tubular muscles, IFMs are asynchronous muscles that utilize a stretch-activation mechanism to contract at a high frequency, 200Hz in *Drosophila*. Contraction of the DLMs deflects thorax and via the wing hinge moves the wings down; simultaneously this stretches the DVMs inducing their contraction, which moves the wings up again, in turn stretching the DLMs. These oscillating muscle movements thus do not require calcium cycling, which would be extremely energy intense at this high contraction rate. Motor neurons fire at lower frequency during flight and Ca^{2+} levels stay elevated for the entire time of flight (Josephson et al., 2000). Furthermore, IFMs possess a very particular composition of contractile proteins and protein isoforms that lead to the characteristic fibrillar morphology. Taken together, *Drosophila* IFMs resemble some features of human heart muscles both in biomechanics and myofibril morphologies. IFMs are dispensable for *Drosophila* life, making them an excellent model to explore gene functions in muscle morphogenesis and function.

1.2.2 Development of *Drosophila* indirect flight muscles

1.2.2.1 Myoblast specification

As stated above (section: 1.2.1.1), AMPs are born as siblings of embryonic founder cells and characterized by persistent Twist expression (Bate et al., 1991). After the first asymmetric division of progenitor cells, FCs inherit Numb from the progenitor, which will repress Notch signaling and thus induce their differentiation. In contrast, Numb is not received in the AMPs, thereby Notch signaling is active and maintains Twist expression (Carmena et al., 1998; Ruiz Gomez and Bate, 1997). Six to seven of these undifferentiated AMPs in each half of the embryonic mesothoracic segment are firmly associated with wing disc primordial (Bate et al., 1991). These AMPs start to proliferate at the end of the second larval instar. One subset of these AMPs is designated to form IFMs, while another subset is going to build direct flight muscles (DFMs). The diversity of the pools of AMPs is defined by the differential expression of *vestigial* (*vg*) and *cut* (*ct*) (Sudarsan et al., 2001). The distal group of AMPs on the wing disc expresses high level of *ct* and low level *vg* is going to give rise DFMs. In comparison, the larger proximal subset of AMPs expresses *vg* and low levels of *ct* will contribute to IFMs. Interestingly, *vg* and *ct* maintain the distinction between the two groups of myoblasts in a mutually repressive manner (Sudarsan et al., 2001).

1.2.2.2 Founder cell specification

A founder cell mechanism (FCs and FCMs) for flight muscle development similar to embryonic somatic muscle specification has been proposed (Bate et al., 1991; Dutta et al., 2004; Rivlin et al., 2000). During the metamorphosis, most of the larval muscles undergo histolysis (~8h after puparium formation (APF)) and adult muscles such as DVMs are formed de novo. In the exceptional case, DLMs forms from the three larval oblique muscles (LOMs) in the mesothorax, which will serve as the templates to which AMPs will fuse. Before fusion, *dumbfounded* (*duf*) expression has been established in a few larger AMPs at the position where DVMs form, and in the DLM templates (Dutta et al., 2004). These *duf* expression cells are set out as FCs and low *duf* expression AMPs are designated as FCMs. Upon ablation of *duf* expressing FCs, myogenesis fails to occur properly, showing the importance of the FC concept (Atreya and Fernandes, 2008).

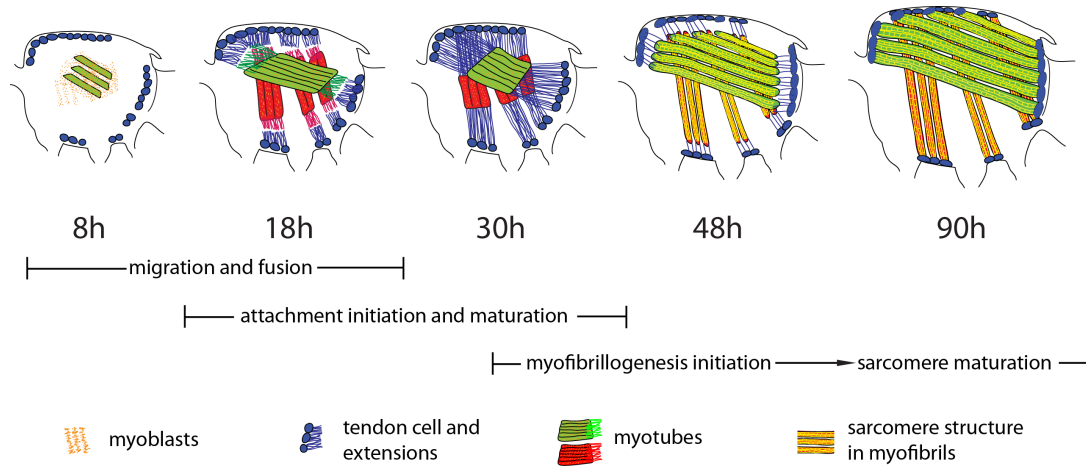


Figure 6: Developmental scheme of IFMs, adapted from Maria Spletter (M. Spletter et al., in preparation). FCMs migrate and start to fuse with DLM templates or DVM FCs from 8h APF. At 18h APF, both tendon cells and myotubes send out extensions to contact each other and establish attachment. Myotubes grow and compact at 30h APF, when myofibrillogenesis starts and connects the myotubes to myofibers. From then on, myofibrillogenesis and sarcomere maturation continues until eclosion.

1.2.2.3 Myoblast fusion

FCMs start to fuse with DVM FCs or DLM templates from ~8h APF until ~24h APF (Figure 6). How the FCMs fuse to DLM templates (larval LOMs) or to DVM FCs is not well studied. However, it is likely that similar principles governing FCMs fusion in the embryo are also used during adult muscle development (Schejter, 2017; Schulman et al., 2015). In analogy to the embryonic somatic FCMs fusion, the multiple cellular events of the fusion process could be divided into four steps: 1) FCMs recognition and adhesion, 2) cytoskeleton rearrangement at the contact site, 3) fusion pore formation at the contact membranes 4) pore expansion and mixing of cellular components.

Four immunoglobulin domain-containing proteins, including Dumbfounded (Duf), Roughest (Rst), Sticks and stones (Sns), and Hibris (Hbs), have been reported to direct the recognition and attraction between FCMs and FCs during embryonic myoblast fusion (Abmayr and Pavlath, 2012). Duf is also expressed in DLMs muscle templates while Sns is expressed in myoblasts. Both are essential for myoblast recognition and establishment of fusion sites between myotubes and myoblasts (Gildor et al., 2012). After recognition, SCAR/Wave and Wiskott-Aldrich Syndrome protein (WASp) activate Arp2/3 actin polymerization machinery at the contact site, which mediates cytoskeletal remodeling and thus promotes the membrane flattening and tight association at the contact site (Kim et al., 2015a; Mukherjee et al., 2011). This results in the formation of pores between the apposed membranes and thereby initiates cytoplasmic components mixing (Kim et al.,

2015b; Mukherjee et al., 2011). Finally, the full cytoplasmic continuity state is reached, which leads to the formation of the myotubes.

1.2.2.4 Muscle tendon attachment

At the same time as myoblasts fuse, tendon precursors are specified. Similar with embryonic attachment sites, tendon cells are specified by *stripe* (*sr*) (Fernandes et al., 1996; Ghazi and VijayRaghavan, 2003). Initially, a few discrete patches of *sr* expression cells within the imaginal discs are associated with overlapping AMP. As the metamorphosis proceeds, these cells developed to tendon cells, which may guide the myotubes towards them, by possibly secreting factors (Fernandes et al., 1996).

At about 18h APF, the growing myotubes and tendon cells send out filopodial extensions to contact each other and establish attachment (Figure 6) (Fernandes et al., 1996; Weitkunat et al., 2014). A recent study shown that Kon-tiki, a single-pass transmembrane protein of the neurexin family, plays essential roles during muscle tendon attachment establishment (Weitkunat et al., 2014). Upon knock down of *kon-tiki* in IFMs, the developing IFMs failed to recognize the tendon extensions and start to round up (Weitkunat et al., 2014). It is also believed that α PS1 and β PS integrin expressed in tendon cells as well as α PS2 and β PS expressed in the myotubes mediate attachments formation by forming heterodimers on the opposed cell membranes (Fernandes et al., 1996). Interestingly, mechanical tension in the muscle tendon system increases as the attachment maturation proceeds (Weitkunat et al., 2014).

Once the contact is made, it becomes smooth and stable. After attachment maturation, myotubes compact (Figure 6). Meanwhile, myofibrillogenesis starts and myotubes are converted to myofibers at 24-30h APF. Myofibrils and sarcomeres mature until 90h APF when the fly encloses and is ready to take off (Reedy and Beall, 1993).

1.3 Myofibrillogenesis

1.3.1 The models of myofibrillogenesis

1.3.1.1 The two historical models

Principles underlying myofibrillogenesis are not well understood. Historically, the pre-myofibril model and the two-state model have been proposed largely based on

knowledge from cell culture (Holtzer et al., 1997; Kontogianni-Konstantopoulos et al., 2009; Ojima et al., 1999; Sanger et al., 2005; 2017).

The first model suggests that the mature myofibrils are formed from the non-muscle myosin (NMHC) containing premyofibrils via by a stepwise exchange of NMHC to Mhc. In the premyofibrils, actin filaments are anchored to Z-bodies by α -actinin and interdigitate with non-muscle myosin II filaments. With the incorporation of titin and other myosin-binding proteins, premyofibrils mature into nascent myofibrils. At this stage, the Z-bodies are better aligned and the non-muscle myosin is exchanging with muscle myosin. Afterwards, the growth and maturation of both thin and thick filaments, as well as the integration of more Z-disc and M-line proteins facilitate the maturation into the final mature stage (Sanger et al., 2005; 2017).

The second model proposed that the formation of I-Z-I bodies and Mhc containing thick filaments is more independent and both components will be assembled at the same time to form the sarcomere structure (Holtzer et al., 1997; Ojima et al., 1999). Before myofibrillogenesis, the globular-actins polymerize to form actin microfilaments and the myosin (including muscle myosin) assembles into bipolar myosin minifilaments through its C-terminal rod tail (Boateng and Goldspink, 2007). The preassembled structures such as minifilament precursors are prerequisite for the initiation of myofibrillogenesis in both models. How myofibrillogenesis is coordinated throughout the myofibers to allow defined myofibril initiation and sarcomere maturation is not well understood/explained by both models.

1.3.1.2 A tension driven model of myofibrillogenesis

It has been shown that contractile activation is necessary for sarcomerogenesis and myofibrillar organization in cultured myotubes and cardiomyocytes, indicating that mechanical tension in the system may be needed for sarcomere assembly (De Deyne, 2000; Simpson et al., 1993). Indeed, *in vivo* studies found that myofibrils are detectable after the force-resistant attachment of myotubes to tendons. Moreover, tension is built up before myofibrillogenesis is initiated (Weitkunat et al., 2014). Importantly, compromising the tension buildup either by laser-cutting or genetic manipulation abolished correct myofibrillogenesis, indicating that tension build up and myofibrillogenesis are coupled processes (Weitkunat et al., 2014; 2017). Based on these findings and previous knowledge, Schnorrer and colleague proposed a novel tension driven model of myofibrillogenesis (Lemke and Schnorrer, 2016). In the initially stage,

Introduction

both ends of myotubes are attached to tendons and tension is built up to orient as well as organize the assembly of the actin and myosin filaments across the whole myotubes (Figure 7A). Subsequently, addition of core sarcomeric components such as muscle myosin and titin to myofibrils increases tension in the system and stimulates the self-organization of periodic immature myofibrils (Figure 7B). At the later stage, more myosin, titin, and α -actinin are added to the periodic immature myofibrils, which leads to the sarcomere maturation as the tension driven self-organization proceeds (Lemke and Schnorrer, 2016; Weitkunat et al., 2014; 2017) (Figure 7C). This newly formatted hypothesis perfectly explains why sarcomerogenesis proceeds simultaneously throughout the entire large muscle fiber.

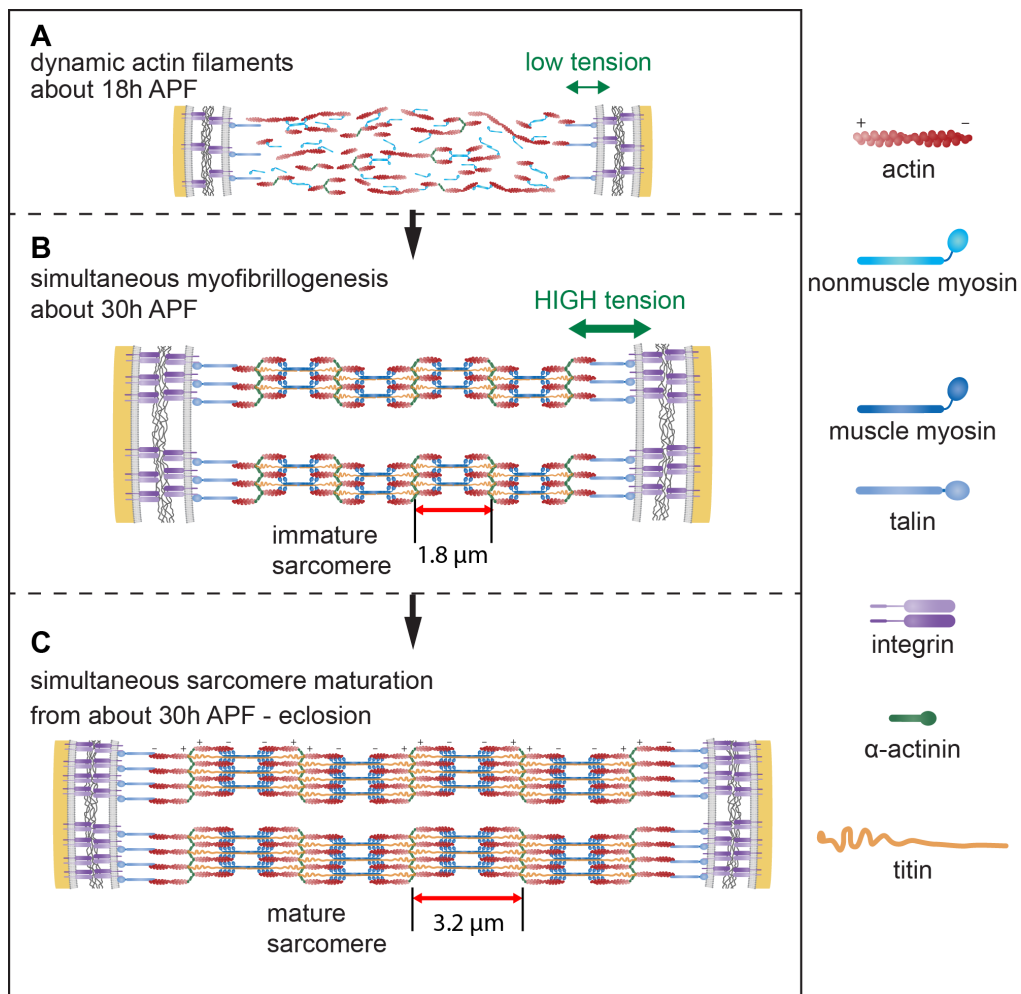


Figure 7: A tension driven myofibrillogenesis model in *Drosophila*, modified from (Lemke and Schnorrer, 2016).

A: Dynamic actin filaments at 18h APF when tension is low. B: At 30h APF, muscle tendon attachment matured and the tension in the tissue is high. B: Immature myofibrils assemble simultaneously to span the entire fiber. Note that at this stage, sarcomere structure is still immature. C: Large amounts sarcomeric structural components are incorporated into the existing immature sarcomeres.

1.3.2 Myofibril and sarcomere maturation in *Drosophila* IFM

After muscle tendon attachment immature myofibrils contain periodic pattern of α -actinin, obscurin, and kettin at 30h APF (Orfanos et al., 2015; Weitkunat et al., 2014) (Figure 7A, 7B). At this stage (30h APF), the immature sarcomeres are about 1.8 μm in length. From then on, the sarcomeres need to mature and continues to grow to 3.2 μm in length and reach a diameter of 1.5 μm in the adult IFM (Orfanos et al., 2015). And interestingly, the number of myofibrils (~2000) in one fiber keeps constant during sarcomere maturation stage, although the length of myofibrils and the volume of fibers grow dramatically during this period (M. Spletter, X. Zhang and F. Schnorrer in preparation). Strikingly, all myofibrils and sarcomeres mature synchronously in both fibrillar IFMs and tubular abdominal muscles (Weitkunat et al., 2014; 2017). A recent study showed that simultaneously growth of sarcomere is realized by incorporating actin monomers and short filaments to the existing thin filaments actively (Shwartz et al., 2016).

Collectively, these data suggest that the sarcomere maturation is involved in synchronously adding of components, including actin filaments and thick filaments, to the immature sarcomere. However, the detailed mechanisms of sarcomere maturation and its transcriptional regulation as well as the coordinated addition of components to all the sarcomeres simultaneously remain to be explored.

1.4 Transcriptional regulation of IFM specification and differentiation

Multiple transcription factors and signaling pathways were shown to be involved in the regulation across all the steps of myogenesis (reviewed in (de Jussineau et al., 2012; Dobi et al., 2015)). Although a number of key identity genes have been extensively studied in *Drosophila* embryonic myogenesis, little progress was made until recently in adult muscles. How transcription factors instruct the formation of myofibrils and sarcomeres, in particular, how they instruct differently in the different muscle types is poorly understood. IFMs, a fibrillar *Drosophila* adult muscle, serve an ideal model to answer these questions.

1.4.1 Mef2 - A general factor

Myocyte Enhancer Factor-2 (Mef2), a MADS domain transcription factor, plays central role during vertebrate muscle differentiation (Black and Olson, 1998). Studies in *Drosophila* myogenesis indicate a function during both embryonic and adult myogenesis (Bour et al., 1995; Bryantsev et al., 2012a; Cripps et al., 1998; Ranganayakulu et al., 1995; Soler et al., 2012). After AMPs specification, FCMs fuse with FCs or DLM templates to initiate differentiation. Before IFMs differentiation, *Mef2* is inhibited by Twist through its target *Holes in muscle (Him)* in myoblasts (Soler and Taylor, 2009). At the onset of myoblast fusion (8h APF), Vg mediated repression of Notch signaling down-regulates Twist, thus its inhibitory effect on *Mef2* is released, which promotes myoblasts and IFMs differentiation (Bernard et al., 2006). Interestingly, upon differentiation, elevated *Mef2* levels activate Vg expression via an 822bp *vestigial* adult muscle enhancer (AME) (Bernard et al., 2009). When *Mef2* is depleted during early stages of fusion, myoblasts migrate to myotubes normally but fail to fuse (Bryantsev et al., 2012a). Furthermore, reduced expression levels of *Mef2* at later stages do not affect the formation of myofibers, but myofibers shown perturbed integrity and severe defects in myofibril organization (Bryantsev et al., 2012a; Soler et al., 2012).

Taken together, *Mef2* plays essential functions at different stages including initial myoblast and myotube differentiation, as well as myofibril maturation (Bryantsev et al., 2012a; Soler et al., 2012). However, *Mef2* binds to pan-muscular-enhancers and many structural gene enhancers, acting as a general factor important for most if not all muscle types. Hence, fibrillar IFMs acquire their specific properties by additional transcriptional regulation.

1.4.2 Identity genes of IFM patterning and differentiation

1.4.2.1 Spalt major (Salm)

In a genome-wide RNAi screen for genes required for adult myogenesis, numerous genes, including transcriptional regulators and structural components, essential for IFMs development have been identified (Schnorrer et al., 2010). Among them, *salm*, a zinc finger transcription factor, has been identified as a fibrillar flight muscle specific selector gene (Schönbauer et al., 2011). *salm* is specifically expressed in IFMs after myoblast fusion, and *salm* known-down switches IFM from fibrillar morphology to tubular morphology. When ectopically expressed in leg muscle, *salm* induces a switch from

Introduction

tubular to fibrillar fate. This indicates that *salm* is not only required but also sufficient to induce fibrillar flight muscle fate in *Drosophila*. The role of Salm is conserved across most insects, thus, *salm* is a myofibril selector gene across 300M years of evolution (Schönbauer et al., 2011). Interestingly, *vg*, another transcription factor expressed in IFM forming myoblasts is required for *salm* expression in myotubes, thus playing a role upstream of *salm* (Bernard et al., 2003; Schönbauer et al., 2011) (Figure 8).

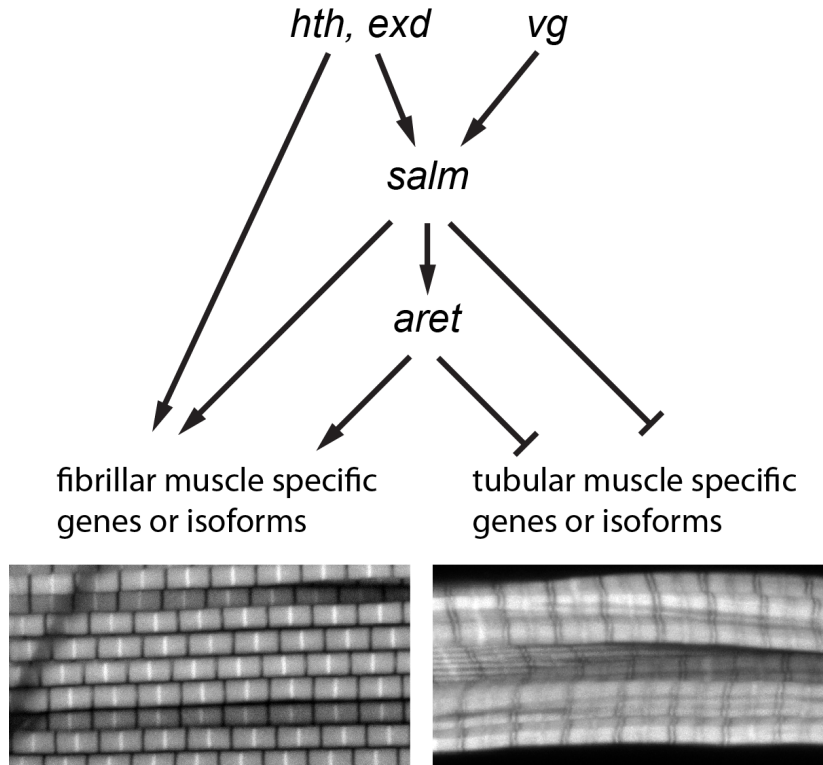


Figure 8: Transcriptional cascade of IFM specification

Exd, *hth*, and *vg* locate upstream of *salm* and positively regulate *salm* expression, while *exd* and *hth* also induce the expression of a subset of fibrillar muscle specific genes. *Salm* is required and sufficient to induce fibrillar muscle fate as well as to repress tubular muscle fate. *Aret*, a downstream target of *salm*, regulates alternative splicing of a subgroup of fibrillar sarcomeric gene isoforms.

1.4.2.2 Extradenticle (Exd) and Homothorax (Hth)

A later study from Cripps lab showed that two homeodomain proteins, Extradenticle (Exd) and Homothorax (Hth), are also required to determine fibrillar muscle fate (Bryantsev et al., 2012a). They are both expressed in developing IFMs and other body wall muscles except the jump muscles from the myoblast stage onwards, while *salm* is mainly expressed in IFM forming myotubes (Bryantsev et al., 2012b; Schönbauer et al., 2011). It has been shown that *exd* and *hth* activate the expression of flight muscle specific structural genes such as *Actin88f* and *flightin (fln)*, and repress the expression of *TpnC41C*, a jump muscle specific structural gene (Bryantsev et al., 2012b). It is worth to

note that *vg*, *exd* and *hth* function genetically upstream of *salm* (Figure 8). Except these selector genes, *E2F*, a cell cycle transcription factor, also has been shown to directly regulate IFMs specific gene transcription (Zappia and Frolov, 2016).

1.4.3 The downstream targets of Salm

During embryonic myogenesis, identity genes have been shown to specify muscle fate by regulating downstream genes essential for different aspects of muscle development (Tixier et al., 2010). The same is true for fibrillar IFM development. By microarray analysis, hundreds of fibrillar muscle specific genes, including the known stretch sensitive genes *fln*, *Myofilin*, *Stretchin-Mlck* (*Strn-mlck*), and *TpnC4*, were found to be significantly downregulated in *salm* knockdown IFMs (Schönbauer et al., 2011). Conversely, many body wall muscle specific proteins such as MP20 and TpnC41 are upregulated (Schönbauer et al., 2011). Notably, *exd* and *hth* need to collaborate with *salm* to regulate many of these IFM specific genes expression (Bryantsev et al., 2012b; Schönbauer et al., 2011). The microarray study also suggested that *salm* not only regulates its targets transcriptionally but also changes alternative splicing. This was systematically investigated by mRNA-seq analysis, which identified a core group of about 700 fibrillar specific gene isoforms are regulated by *salm* (Spletter et al., 2015). Thus, *salm* regulates IFM fate determination and differentiation by differential transcription and splicing.

Although some targets of Salm have been found, it is unclear if Salm binds to the enhancers of these genes directly and how it collaborates with Mef2. Moreover, as discussed above (1.3.2 and 1.3.3) myofibrillogenesis occurs in two stages, namely myofibrillogenesis initiation and sarcomere maturation, during *Drosophila* IFM development. Does Salm regulate these two stages directly or by its downstream transcription factors? To gain insights about the underlying mechanism of transcriptional regulation, it is necessary to identify direct targets of Salm in IFMs. However, it is difficult to collect enough fresh and pure myotube material and to obtain high quality antibodies for genome-wide analysis, such as ChIP-seq studies in a tissue and stage specific manners. Thus, establish precise genome editing method to tag the endogenous *salm* locus is important, which enable one to use commercially available high quality antibodies.

1.4.4 Splicing regulation during IFM development

aret (also called *Bruno*), a *Drosophila* orthologue of vertebrate *CELF* (*CUG-binding protein and ETR-3-like factor*) RNA-binding protein, has been suggested as a potential *salm* target by microarray data analysis. This has been confirmed by two recent investigations which found that *aret* is directly involved in the regulation of IFMs specific protein isoforms splicing such as Strn-Mlck and Tpn I (Oas et al., 2014; Spletter et al., 2015). Knockdown of *aret* in IFMs disrupt the myofibril maturation process largely after 48h APF. At 72h APF, the sarcomere length is significantly shorter compared to wild type. The *aret* mutant IFM fibers hypercontract resulting muscle atrophy at the adult stage (Spletter et al., 2015). These results showed that post-transcriptional regulation also plays an essential role during IFM sarcomere maturation. Interestingly, it has been shown that in vertebrates, both transcriptional regulation and alternative splicing are extensively involved in muscle fiber specification (Spletter and Schnorrer, 2014). However, how the defective sarcomere maturation caused by misregulated transcriptional program is not clear.

1.5 IFMs as a model to study human muscle hypercontraction

1.5.1 Conservation of muscle biology from fly to human

Compared with the single-fiber embryonic muscles, *Drosophila* adult muscles are composed of multiple fibers, which resemble vertebrate skeletal muscles and make it a more attractive model for vertebrate muscle development. Despite the overall difference between vertebrate skeletal muscle and *Drosophila* flight muscle, many developmental and functional aspects are conserved.

First, patterning and regulation of muscle development in both systems are particularly well conserved (summarized by Taylor (Taylor, 2006)).

Second, the functional core complexes (actin/troponin-tropomyosin, titin and myosin) and their arrangement in both *Drosophila* adult muscles and human muscles are very similar. And not surprisingly, they share a common contracting mechanisms: the filament sliding and myosin swinging cross-bridge model (Hooper and Thuma, 2005; Hooper et al., 2008).

Moreover, *Drosophila* adult flight muscles are dispensable for life, which allows the investigation of null mutations. With this system, using the large body of genome-wide resources, it is easier to genetically manipulate them, compared to vertebrate models. Biological research in *Drosophila* showed its large potential in unraveling vertebrate development in normal and pathological conditions (Ugur et al., 2016; Wangler et al., 2015). Thus, studying *Drosophila* adult muscle development will likely impact vertebrate muscle biology.

1.5.2 Model for muscle hypercontraction

Distal arthrogryposis (DA) syndromes are rare and complicated congenital diseases characterized by uncontrolled distal limb joint contractures. They are often associated with congenital myopathies (Kimber, 2015). Interestingly, several DA subtypes have been identified as a consequence of mutations in genes encoding sarcomeric proteins such as *β -tropomyosin (TPM2)*, *fast troponin I (TNNI2)*, *fetal myosin heavy chain (MYH8)*, *fast troponin T (TNNT3)*, *embryonic myosin heavy chain (MYH3)*, and *slow skeletal muscle myosin binding protein C (MYBPCI)* (Gurnett et al., 2010; Kimber et al., 2006; Sung et al., 2003; Tajsharghi et al., 2007; Toydemir et al., 2006a; 2006b). Strikingly, these DA-associated mutations in *TPM2*, *TNNI2*, and *TNNT3* all display elevated muscle contractility and produce muscle hypercontraction (Robinson et al., 2007). These results explain why myopathies are usually observed in DA patients. It is worth to note that a *MYBPCI* mutation also has been detected in another subtype of arthrogryposis called lethal congenital contractural syndrome (LCCS) (Markus et al., 2012). Thus, it appears that the multiple contracture phenotype results from the uncontrolled force caused by misregulated actinomyosin machinery during muscle development in humans.

A series of mutations in sarcomeric actinomyosin components are known in flies. Some of these are associated with muscle hypercontraction, a phenotype characterized by normally developed IFMs which undergo excessive contractions causing muscle damage, muscle detachment, and subsequent muscle atrophy (Nongthomba et al., 2003). Mutations in IFMs specific isoforms of *Tpn I* or *Tpn T* can lead to muscle hypercontraction and subsequently muscle detachment or muscle fiber tearing in the middle of the fibers (Nongthomba et al., 2007; 2004). At the sarcomere level, the IFMs display Z disc and M-line disorganization and so-called zebra-bodies, which are actin accumulations. Furthermore, mutations in the fly single muscle myosin heavy chain *Mhc*

gene, which reside in the ATPase domain (*Mhc*⁶ and *Mhc*¹³) also have been reported to cause IFMs hypercontraction (Montana and Littleton, 2004).

In summary, as the core components of the sarcomere are highly conserved from *Drosophila* to human, studying the function of these proteins during IFM development may gain mechanistic insights into the pathologies of human muscle diseases.

1.6 Genetic *Drosophila* resources

1.6.1 RNAi libraries, Fosmid libraries, and MiMiC collections

In the past genomic era, remarkable efforts have been put into the development of genome-wide approaches and resources. This led to the generation of the first *Drosophila* genome-wide transgenic RNAi library (Dietzl et al., 2007). Meanwhile, additional RNAi libraries including NIG-FLY and TRiP were created (summarized in (Kaya-Çopur and Schnorrer, 2016)). These resources largely expanded the systematic assessment of gene function. One such genome-wide tissue-specific screen was performed in *Drosophila* muscle and identified numerous novel players important for muscle development and function (Schnorrer et al., 2010). Among these, *Salm* has been identified as a conserved identity gene for IFM development in insects (Schönbauer et al., 2012). Another well-established example is *Aret*, which also has been confirmed an important function during IFM differentiation (Spletter et al., 2015). Apart from these, several interesting candidate genes, including *CG11617*, *CG17912*, and *CG32121*, have been found to be specifically expressed in *Drosophila* body muscles (Sarov et al., 2016). Their functions are still waiting to be confirmed with mutants.

The next step after the discovery of a candidate gene is a mechanistic characterization of its function. Usually, the first step is to define the particular expression and subcellular localization of the candidate protein. However, due to lack of specific antibodies or live visualization probes, it is time consuming to make antibodies or tag the gene of interest with traditional methods. Thus, two collections of protein tag libraries have been established in flies: a Minos Mediated Integration Cassette (MiMIC) collection and a Fosmid collection (Nagarkar-Jaiswal et al., 2015; Sarov et al., 2016; Venken et al., 2011). MiMIC applied a genome-wide mutagenesis by a minos transposon mediated insertion of a cassette carrying a dominant marker and STOP cassette flanked by two integrase attP sites. Upon insertion, within an intron the termination signal followed the

splicing acceptor sequence in the cassette terminates transcription resulting in a loss of function allele. By Φ C31 integrase mediated cassette exchange, the termination cassette can be exchanged with any DNA of interest. This smart design allows versatile genome-manipulation such as creating a protein-trap line for studying protein localization. One important limitation for the MiMIC system is that the insertion position is random, thus the insertion may affect protein function and not all genes of the genome are accessible.

To complement this limitation, a strategy with high gene coverage and defined tag position would be preferable. The fosmid library combines recombineering approaches to tag a large *Drosophila* genomic DNA fragment in bacteria, with Φ C31 integrase mediated transgenesis into a defined locus to generate transgenic *Drosophila* lines (Ejsmont et al., 2009). With this approach, a comprehensive genome-wide fosmid constructs library including about 10,000 C-terminal tagged proteins or protein isoforms clones, has been created. These clones were used to generate more than 800 transgenic lines and about 200 tagged proteins have been tested for their expression and cellular localization pattern *in vivo* (Sarov et al., 2016). The results showed a impressive potential for using fosmid lines to investigate protein expression and localization pattern, as well as protein-protein interaction, and protein dynamics with live imaging.

1.6.2 Precise genome editing methods: TALENs and CRISPR

The development of precise genome-editing methods significantly boosted the study of reverse genetics in recent years by allowing locus specific gene targeting. In order to fully explore the function of novel genes, it is necessary to modify the endogenous gene locus. However, it is relatively tedious to engineer a specific gene with traditional homologous recombination approaches in flies. The recent development of Transcription Activator-like Effector Nucleases (TALENs) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) technologies in other model organisms provide potential ways to improve the gene targeting in *Drosophila*.

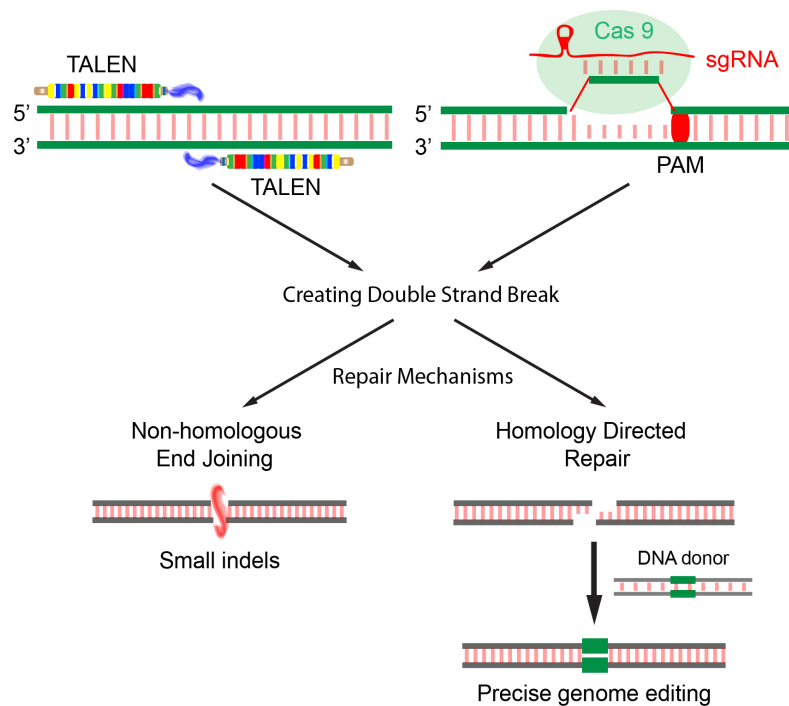


Figure 9: Schematic representation of TALEN or CRISPR/Cas mediated genome engineering.

TALEN or Cas9-sgRNA complex bind to DNA and induce DSBs at the designed position. For Cas9 induced cut, a PAM sequence is necessary to allow Cas9-sgRNA complex to unwind the genome. The generated DSBs will be repaired either by error-prone NHEJ pathway or by HDR in the presence of homology donor sequence. The latter enables precise genome editing.

TALENs originate from naturally occurring transcription factors present in plant pathogens called TALEs. They are built of 12–27 repeats, each typically containing 34 amino acids. Each repeat binds to a specific base pair of DNA, which is determined by a di-residue within the repeat. The C-terminal transcriptional activator of TALEs has been replaced with a FokI nuclease domain, converting it to a sequence specific nuclease (Barrangou, 2013; Joung and Sander, 2012). Since the FokI nuclease is active as a dimer, TALENs are designed in pairs binding to opposite DNA strands (Figure 9). The double strand breaks (DSBs) generated by TALENs are repaired by non-homologous end joining (NHEJ) or by homology directed repair (HDR) in the presence of homologous sequences (Figure 9). NHEJ usually introduces small insertions or deletions and thus causes mutations, while HDR is a precise repair, which can be used for precise gene tagging (Joung and Sander, 2012). TALENs had been applied to several organisms including *Arabidopsis*, *zebrafish*, and *mice*, to mutate or tag genes but not yet to *Drosophila* when this work was initiated (Cermak et al., 2011; Sung et al., 2013; Zu et al., 2013).

Most recently, the application of CRISPR/Cas9 system for gene targeting is booming, largely promoted by its simple design and high efficiency to introduce DSBs.

CRISPR/Cas9 is a RNA-guided DNA endonuclease from *Streptococcus pyogenes*. It can form a complex with single strand guide RNA that is designed to bind to specific genome sequences. A Protospacer Adjacent Motif (PAM) located at the 3' end of the targeting sequence is required for Cas9-sgRNA complex to interrogate into the DNA (Sternberg et al., 2014). Once the complex binds to the genome, it cuts the targeting region and induces DSBs that can be repaired similarly as the TALENs generated DSBs. Thus CRISPR can also be utilized to mutate or tag a gene (Figure 9). CRISPR/Cas has been successfully used in yeast, eukaryotic cells, mice, as well as zebrafish and was successfully applied to *Drosophila* in the course of this work (Bassett et al., 2013; Cong et al., 2013; Dicarolo et al., 2013; Gratz et al., 2013; Hwang et al., 2013; Wang et al., 2013).

1.7 Aims of the thesis

In this work, I aimed to establish precise genome-editing methods in *Drosophila melanogaster* and then applied them to dissect the function of genes involved in *Drosophila* IFM development.

First, I planned to establish and use TALEN to mutate candidate genes (*salr*, *CG11617* and *CG32121*) from our RNAi screen.

Second, I aimed to establish CRISPR/Cas and employ it to mutate *CG32121* (which we called *Fallen Angel*, *fall*), to tag *fall*, and to generate different gene variants for structural function study. I also planned to apply CRISPR to tag *salm* at its N-terminus and C-terminus with different tags for the ChIP-seq experiment in the future.

Third, I wanted to characterize the phenotype of *fall* mutants, the expression of Fall, the molecular function of Fall protein domains, and its role in IFM during development especially during sarcomerogenesis.

Introduction

2 Discussion and Outlook

2.1 Establishment of precise genome editing methods in *Drosophila*

The emergence of TALENs and CRISPR mediated genome modifying methods revolutionized our view of genome editing. Shortly after the pioneering publications, these methods have been adapted to several model organisms and showed their potential for human gene therapy (Doudna and Charpentier, 2014; Hsu et al., 2014; Mendell and Rodino-Klapac, 2016; Themeli et al., 2015).

2.1.1 Applications of TALENs in *Drosophila*

We have successfully established TALENs induced mutagenesis in *Drosophila* at the same time as other labs (Katsuyama et al., 2013; Zhang et al., 2014a). Null mutants of our candidate genes from the RNAi screen, including *CG32121*, *CG11617*, and *spalt related*, have been generated. Surprisingly, *CG11617* and *spalt related* null alleles did not give any obvious phenotype in IFM (X. Zhang and F. Schnorrer unpublished data), while *fall* mutant displayed a severe defect during IFM development. In our hand, we used TALENs to generate specific gene knockout with a relative high efficiency (up to ~10% mutagenesis per fertile G0 cross) (Zhang et al., 2014a).

However, we have not been successful in integrating a GFP tag by HDR despite testing several conditions (unpublished results). This was likely due to the limitation of our PCR-based detection method of the rare HDR event. In other studies, only a few genes with visible phenotypes, which enable a very simple identification of the targeted animals, have been targeted by HDR (Katsuyama et al., 2013; Liu et al., 2012). These observations indicate that the HDR efficacy would need further optimization in order to effectively tag or manipulate genes systematically. Moreover, the tedious assembly of TALEN pairs is an obvious disadvantage compared with CRISPR/Cas (Cermak et al., 2011).

2.1.2 CRISPR/Cas9

2.1.2.1 General applications of CRISPR/Cas9 in *Drosophila*

Compared with TALENs, the CRISPR/Cas9 system only requires one single strand guide RNA (sgRNA) and the presence of the Cas9 nuclease to cut genomic DNA at a defined position (Cong et al., 2013). Several algorithms provided online facilitate the design of sgRNAs including the one that we have used from the Feng Zhang lab (<http://crispr.mit.edu/>). In *Drosophila*, multiple strategies have been designed for customized applications based on CRISPR/Cas9 systems (Bassett et al., 2013; Gratz et al., 2013; Zhang et al., 2014b). The simplest application, generation of gene knockout, has been implemented in several publications with high efficiency. The CRISPR/Cas9 generated indels can be easily identified with various methods, including the T7 endonuclease assay that we applied, High-Resolution Melt Analysis (HRMA), and surveyor nuclease assays (Bassett et al., 2013; Cong et al., 2013; Zhang et al., 2014b). By applying a selection marker, the identification of germline genome editing events is strongly simplified. Another application is locus specific insertion of DNA elements such as tag sequences, an attP sequence, or an FRT site, which enables sophisticated subsequent gene modification (Gratz et al., 2014; Zhang et al., 2014b). These insertions can also be identified by T7 endonuclease or surveyor nuclease assays. However, the frequency for DNA insertion events is much lower compared to small deletions and thus demand more work for identification. A dominant selection marker is preferred to simplify the identification of targeting event.

2.1.2.2 Combination of CRISPR/Cas with RMCE

In many scenarios, researchers want to study their favorite genes in detail, requiring various alleles. This often requires to knockout the gene, and subsequent characterization of its expression and localization pattern. Frequently, a conditional allele would be very insightful. To provide a solution to this problem, we have combined CRISPR with the MiMIC technologies to generate a versatile two-step strategy that enables multiple aforementioned applications (Zhang et al., 2014b) (Figure 10). First, we apply CRISPR to insert a modified MiMIC cassette bearing a dsRed selection marker, which largely simplified the identification of successful insertion events, and STOP cassette into a specific targeting site (Figure 10B). In the second step, recombinase-mediated cassette exchange (RMCE) by the Φ C31 integrase allows to insert any DNA sequences to the

desired locus (Figure 10C). The successful exchange event can be easily found by screening the loss of dsRed marker. We have successfully used this method to tag *salm* at both its N and C-terminus with various tags (Zhang et al., 2014b).

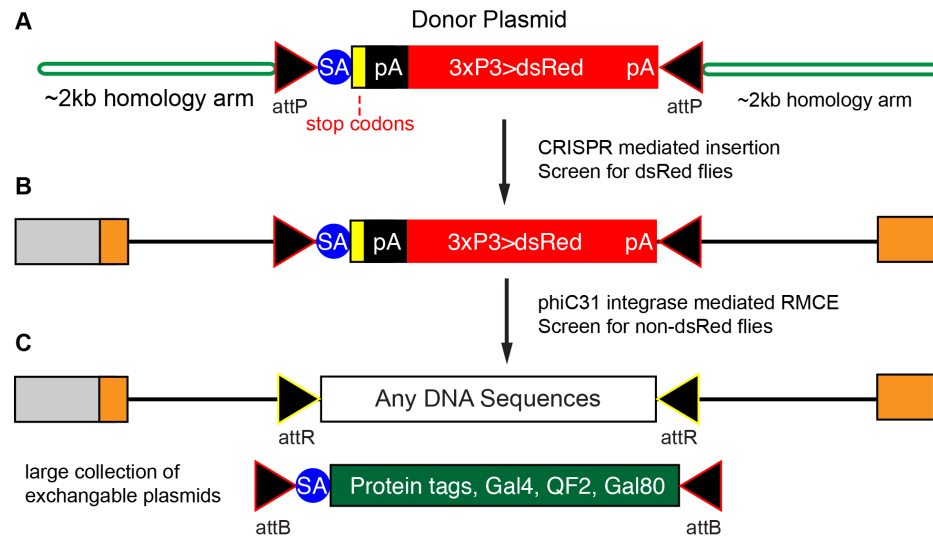


Figure 10: CRISPR-RMCE based strategy for genome editing.

A: A selection cassette bearing a 3P3-dsRed eye marker and two attP sites. B: CRISPR mediated cassette insertion event can be easily identified by dsRed marker screening in flies. C: Φ C31 integrase mediates cassette exchange; successful exchange is identified by non-dsRed flies screening. In this step, any customized DNA sequence can be inserted into the locus of interest.

We have been able to delete mostly the entire *CG32121* gene, which we then called *fallen angel* (*fall*), and have exchanged back different truncated variants, which allow subsequent structural function characterization, into the *fall-dsRed* locus by RMCE. In our lab, we have targeted several loci with the CRISPR-RMCE strategy and obtained an efficiency up to 38% per fertile G0 cross (X. Zhang and F. Schnorrer in preparation). Moreover, a 18kb genomic region has been successfully deleted with a comparable efficiency (M. Spletter and F. Schnorrer unpublished data), showing that CRISPR-RMCE is a versatile tool for genetic manipulation in *Drosophila*.

However, we are aware of the fact that leftover sequences such as attR and FRT sites may interfere with endogenous transcription or splicing regulation and thereby disrupt normal transcription or alternative splicing. From our experience, choosing the targeting site wisely can circumvent this problem. However, one could also use the so-called scarless strategy by putting the selection marker in between two *piggyBac* transposon sites, which can be removed later with almost no leftover sequences (Fei Xie, 2014). In conclusion, CRISPR/Cas9 is highly efficient and flexible to create knockout and knockin

alleles. In combination with RMCE, it enables efficient gene targeting and structural function studies of the endogenous locus of *Drosophila* genes in the future.

2.2 *Fallen angel (fall)*, a novel BTB-ZF transcription factor, is essential for sarcomere growth and maturation

2.2.1 The *fall* loss-of-function phenotype

2.2.1.1 *fall* loss-of-function flies are viable but flightless

With TALENs and CRISPR mutations in *CG32121*, which we called *fall*[1] and *fall*[2] were generated. In *fall*[1], the sequences between the second intron and the end of the second coding exon were deleted, while in *fall*[2] except the first small exon, all other parts of the gene have been deleted. Both *fall*[1] and *fall*[2] are homozygous viable and flightless. In the adult flies, flight muscles undergo atrophy. Detailed analysis showed that this muscle defect arises at around 56h APF during sarcomere growth. In the *fall* mutants, the myofibrils contain shorter sarcomeres. These observations indicate that Fall is dispensable for the viability but essential for sarcomere growth and maturation during flight muscle development.

2.2.1.2 Sarcomere maturation defect and hypercontraction in *fall* mutants

In both *fall*[1] and *fall*[2] homozygous alleles or over a large *fall* deficiency, IFMs undergo muscle atrophy during development with most IFMs degenerating until adult stage. In *fall*[1], tendon cells rupture happens at 56h APF. The MTJ in *fall*[1] appears morphologically normal and β -PS integrin localization was not affected prior to rupture. Furthermore, MTJ reinforcement could not prevent the rupture in tendon cells. However, loss of force in the *Mhc*[10] (Cripps et al., 1994) allele rescued the tendon rupture phenotype. These data indicate that *fall*[1] IFMs produce higher force than the tendon cell can withstand, resulting in their rupture.

In *fall*[2], muscle tearing in the middle of the myofiber, a characteristic of muscle hypercontraction, is the predominant phenotype. The sarcomere morphology in *fall*[2] is affected more severely compared with *fall*[1]. Interestingly, the muscle fiber tearing phenotype in *fall*[2] is also rescued by *Mhc*[10].

Taken together, in both alleles, sarcomerogenesis is disrupted at the sarcomere maturation stage. This induces higher force in the muscle-tendon system that likely exceeds normal levels, leading to hypercontracted muscle and muscle or tendon ruptures. The force in *fall[1]* muscle fibers might be higher since the myofibrils retained a better structure compared with *fall[2]* before rupture. This may explain why the rupture occurs within tendon cells in *fall[1]* while the muscle fibers rupture in *fall[2]*.

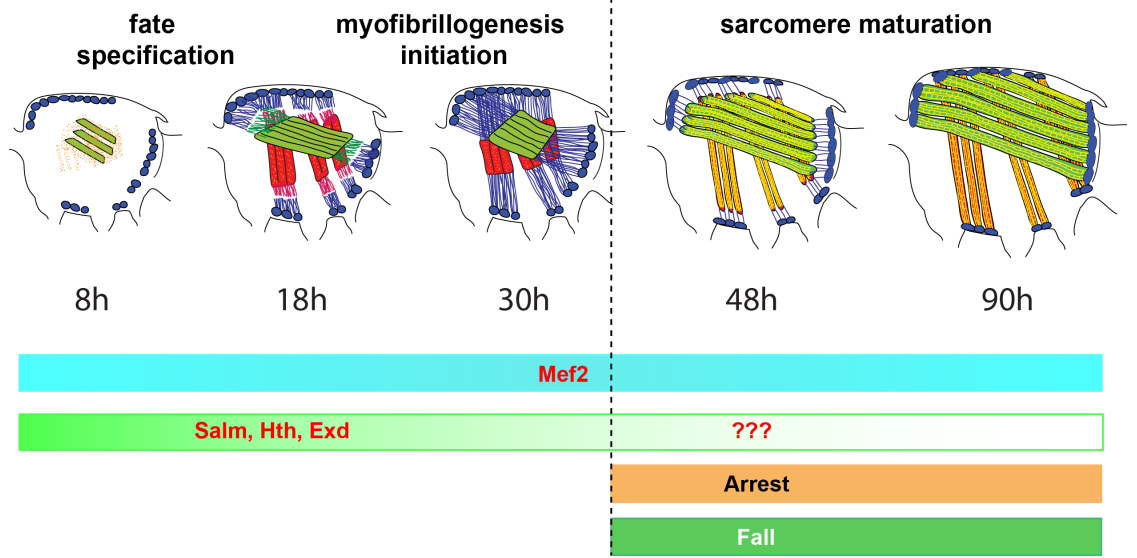


Figure 11: Transcriptional regulation of sarcomere maturation. Mef2 is essential for all stages of IFM development. Hth, Exd, and Salm are known specify IFM fate. Aret is a Salm target and involved in sarcomere maturation. Fall is a novel transcription factor that regulates sarcomeric structural genes expression during maturation stage.

2.2.2 Fall regulates the transcription of sarcomeric genes

2.2.2.1 Fall is a transcriptional regulator

By comparing the mRNA profile of *wild type* with *fall[1]* mutant IFMs at 48h APF, hundreds of misregulated genes have been identified. Moreover, proteomics data showed high correlation with these RNA-seq results, confirming the RNA sequencing data. In contrast, the *DEXSeq* analysis did not show an obvious splicing defect in *fall[1]*. Thus, it is likely that Fall binds to DNA directly and regulates transcription (Figure 11), similar as many other BTB-ZF proteins do. Notably, clustering analysis of RNA-seq discovered that misregulated genes in *fall[1]* are enriched in a particular gene clusters that is upregulated during the sarcomere maturation phase (cluster 22) (M. Spletter, X. Zhang and F. Schnorrer in preparation) (Figure 11). Genes in this cluster, regulated by Fall, include *Zasp52*, *Zasp66*, *Unc-89*, *starvin (stv)*, *Formin homology 2 domain containing (Fhos)*, *sarcomere length short (sals)*, *fln*, IFM specific isoform *Strn-Mlck*, and the titin

homolog *sallimus* (*sls*), which are abundant and essential components for sarcomere maturation during IFM development (Figure 11). Strikingly, all of them are downregulated in *fall[1]* compared with *wild type* at 48h APF, suggesting a defect during sarcomere maturation in *fall* mutants. This is consistent with normal initiation of myofibrillogenesis in *fall* mutants. However, these immature sarcomeres failed to mature properly, resulting in hypercontraction and muscle atrophy. These data let us to propose that Fall is an essential transcription factor that specifically involved in sarcomere maturation stage (Figure 11).

2.2.2.2 Sarcomere maturation defect and hypercontraction might be caused by misregulation of Fall targets

Integrated analysis of RNA-seq and proteomics data showed that the tubular muscle specific *TpnC* (*TpnC47D* and *TpnC25D*) and ubiquitously expressed *TpnC73F* (Herranz et al., 2004) are upregulated while the IFM-specific *TpnC* (*TpnC4*) (Feng et al., 2003) is downregulated in IFMs of *fall* mutants. The misregulation of *TpnC* isoforms may lead to altered troponin/tropomyosin complex dynamics or Ca^{2+} regulated muscle contraction, and thereby influence actinomyosin activity resulting in hypercontraction. In addition, the expression of *Strn-Mlck*, a potential myosin light chain kinase, is almost lost in both *fall[1]* and *fall[2]*. It has been shown that this IFM specific isoform of Strn-Mlck localizes to thick filament and causes IFM hypercontraction upon RNAi knockdown (Spletter et al., 2015). Thus, *fall* mutation induced hypercontraction might be caused by the misregulation of its potential target genes such as *TpnC4* and IFM specific *Strn-Mlck*. Furthermore, in *wild type*, expression of many essential structural genes is induced at 48h APF, in order to produce large amounts of proteins for correct and fast sarcomere growth and maturation. In *fall* mutant, this induction failed to occur, which may lead to sarcomere growth arrest or uncoordinated sarcomere assembly and thus changing the tension balance in muscle tendon system. Taken together, the disrupted actinomyosin activity as well as perturbed sarcomere growth is likely caused by the misregulation of Fall targets and leads to muscle hypercontraction during IFM development.

Interestingly, flies bearing mutations in the troponin complex (*TpnI*, *TpnT*, and *TpnC*) and several dominant myosin alleles such as *Mhc*⁶ and *Mhc*¹³ displayed a specific hypercontraction (hypercontracture in vertebrate) phenotype in IFM during development (see introduction section 1.5), which is characterized by hypercontracted muscle fibers, shortened sarcomeres, and muscle rupture, resulting in muscle atrophy. Moreover, a

more recent publication suggests that the Ca^{2+} -dependent spontaneous contraction is crucial for sarcomerogenesis in *Drosophila* skeletal muscle (Weitkunat et al., 2017). These results and our own data suggest that tightly regulated actinomyosin activity is essential for correct sarcomerogenesis and disruption of actinomyosin activity may lead to muscle hypercontraction during development.

2.2.3 Fall expression and localization pattern

2.2.3.1 Body muscle specific Fall expression and function

The C-terminally tagged Fall-GFP Fosmid line completely rescued the muscle morphology and flight phenotype of *fall[1]* and *fall[2]*. This showed the specificity of the Fall loss-of-function phenotype and validated that Fall is required to prevent IFM atrophy. It also suggests that C-terminal tagging of Fall produces a functional Fall protein, which is also supported by the *fall-CDS-HA* insertion in the endogenous *fall* locus that restored the Fall function. However, a small N-terminally tag prevented the entry of Fall into the nucleus in S2 cells (unpublished observation). Together, this indicates that the C-terminal tagging retains Fall function while N-terminal tagging may interfere with a potential nuclear localization signal.

Both Fall-GFP Fosmid and Fall-HA endogenous tag showed that Fall is expressed in nuclei of all body muscles of the fly but not in visceral muscle. However, muscle atrophy, resulting from *fall* mutation only manifested in flight muscles (DVMs and DLMs) but not jump, leg, and abdominal muscles. *fall* null mutants are fully viable, even when aged for weeks, suggesting that Fall is not essential for muscle function in general. In addition, flight muscle specific expression of Fall with *Actin88f>GAL4* (high expression starts at 18h APF during IFM development) in the *fall* mutation background is sufficient to completely restore IFMs morphology. These data strongly suggest that Fall has a specific function in IFM development but not in other body muscles. Interestingly, rescue of the *fall* phenotype with *Fln>GAL4* (high expression starts after 48h APF during IFM development) was not successful, indicating an early requirement of Fall expression for sarcomere maturation.

2.2.3.2 Fall localises in nuclear body pattern in IFM

Fall expression in IFMs is detectable from 34h APF and reaches high levels at 48h APF. The protein localizes to the nuclei, where it displays multiple discrete small dots or

speckles, some of which have been characterized as a biophysical liquid-liquid phase separation (Zhu and Brangwynne, 2015). This localization pattern is distinct from PML bodies, P bodies, and Cajal bodies but resembles nuclear speckles, paraspeckles, and insulator bodies pattern (Gómez-Díaz and Corces, 2014; Mao et al., 2011). Nuclear speckles and paraspeckles have been found to play roles in pre-mRNA splicing and A-to-I editing (Mao et al., 2011). Previous work found that upon transcriptional inhibition by applying actinomycin D to cell culture, both nuclear speckles and paraspeckles fuse to larger speckles (Shav-Tal et al., 2005). However, the Fall pattern does not change after actinomycin D treatment (data not shown). In addition, immunostaining experiments confirmed that Fall did not show any overlap with SC-35, a nuclear speckle marker (data not shown). These results are consistent with no significant splicing defect detected in *fall* mutants.

Interestingly, the appearance of the Fall speckles is also similar to insulator bodies, another type of liquid-liquid phase separation which has been proposed to regulate 3D genome organization (Gómez-Díaz and Corces, 2014). Upon osmotic stress, insulator bodies fuse together and form large bodies localized to the nucleus periphery (Schoborg et al., 2013). Indeed, Fall showed a similar pattern when NaCl or sucrose induced osmotic stress was applied. However, we only found partially association of Fall with CP190, a marker of insulator bodies in *Drosophila*, upon osmotic stress. Thus, Fall may form a novel class of insulator proteins or represents a new group of proteins capable of forming liquid-liquid phase separation that influences transcription. Noteworthy, we observed that bona fide insulator body proteins do also not overlap entirely. In developing IFM, CP190 and BEAF-32 bodies do not completely overlap while CP190 and Mod(mdg4)67.2 showed an identical localization pattern.

2.2.3.3 Structural function analysis of Fall

Fall protein consists of a N-terminal BTB domain and two C-terminal zinc finger motifs. We have investigated their function separately by replacing the endogenous gene with Fall protein variants, including Fall-2HA, Fall-ΔBTB-2HA, and Fall-ΔZFs-2HA, at the endogenous locus. Their expression has been confirmed by HA staining. In *fall-2HA* flies, Fall protein showed a identical localization pattern with Fall-GFP and the IFM developed normally and the flies fly. Interestingly, flies lacking either BTB or ZFs displayed a similar phenotype to the *fall* null allele; they are homozygous viable, flightless, and show the same muscle atrophy as the *fall* null. By immunostaining, we

have confirmed that deletion of either of the two domains results in hypercontracted muscle fibers with significantly shortened sarcomeres, suggesting that both domains are required for normal sarcomere maturation during IFM development. BTB deletion leads to a nuclear localization defect of Fall which is prominent at 48h APF. In contrast to wild type or Fall- Δ ZF-HA, the Fall- Δ BTB-2HA protein is almost exclusively located to the cytoplasm rather than nuclei. At 72h APF, slight enrichment of Fall- Δ BTB-2HA in nuclei has been seen. Detailed characterization revealed that the Fall- Δ BTB-2HA failed to form liquid-liquid phase separation after osmotic stress treatment, while Fall- Δ ZF-HA behaves similarly to Fall protein. These results strongly suggest that both BTB and ZFs are important for correct Fall function in IFMs. However, it is unclear if the nuclear body formation is coupled to the physiological function of Fall, since it is unknown which part of *fall* is essential for entering nuclear bodies but dispensable for normal localization to nuclei.

2.2.4 BTB-zinc finger proteins and conservation

2.2.4.1 BTB-ZFs are conserved across organisms

BTB-ZF (Broad-complex, Tramtrack and Bric-a-brac Zinc Finger) proteins are a large group of transcriptional regulators characterized by a N-terminal localized BTB domain and a series of C-terminally localized zinc finger motifs. They have been found in *Drosophila* proteins Broad-complex, Tramtrack, and Bric-a-brac initially and named after them (Siggs and Beutler, 2012). Hundreds of proteins belonging to this family have been found from virus, insects to humans. Roles for BTB-ZF proteins in several developmental and pathological processes including neuronal patterning, axon branching, leukemic translocation, and tumorigenesis are established (Siggs and Beutler, 2012). The N-terminally localized BTB domain is frequently involved in dimerization or multimerization, while the zinc finger domains often bind to DNA and determine sequence specificity. Apart from promoting dimerization, BTB domain also recruits other transcription regulators such as NcoR (nuclear receptor corepressor) and SMRT (silencing mediator for retinoid and thyroid hormone receptor) or chromatin remodellers CTCF and BEAF-32 to modify the chromatin status (Dhordain et al., 1997; Gómez-Díaz and Corces, 2014; Huynh and Bardwell, 1998), and thereby repress or activate the expression of their targets. Noteworthy, in *Drosophila*, a group of BTB-ZF proteins, including CP190, Su(Hw), and Mdg(mdg4)67.2, interact with each other and are capable of forming the insulator bodies in order to partition chromatin domains, and thus regulate

transcription (Gómez-Díaz and Corces, 2014). However, despite of a few established roles of BTB-ZF proteins, the functions and downstream targets of many BTB-ZF proteins are unknown.

2.2.4.2 Nuclear body formation: a common feature of BTB-ZF proteins?

In *Drosophila*, several BTB-ZFs proteins, including Su(Hw), Mod(mdg4)67.2, CP190, and Fall potentially localize to different liquid-liquid phase separation compartments. Surprisingly, in vertebrates, a few well-characterized proteins such as PLZF and Bcl-6 have also been found to be able to form nuclear bodies, which are also enriched for their binding partners SMRT and NcoR (Dhordain et al., 1997; Huynh and Bardwell, 1998; Koken et al., 1997; Reid et al., 1995). Notably, they exert their function by deacetylating histones, and thereby regulate chromatin states (Downes et al., 2000). These results indicate that nuclear body formation may represent a general character of BTB-ZFs, and nuclear bodies provide a restricted compartment for particular BTB-ZFs to function by specific chromatin modification.

During the past decades, several nuclear bodies have been identified and characterized (Mao et al., 2011). However, their biophysical properties have been explored only recently (Zhu and Brangwynne, 2015). Biophysical investigations have suggested that the nuclear bodies or non-membrane bound compartments are organized by long non-coding RNAs or protein polymers and are liquid-liquid phase separations (Hyman et al., 2014; Mitrea and Kriwacki, 2016). The composition and function of nuclear bodies is still intriguing. Proteins enriched for tandem repeats of individual amino acids or repetitive amino acid motifs such as polyasparagine (PolyN) and polyglutamine (PolyQ) are often found to locate to liquid-liquid phase separations (Bergeron-Sandoval et al., 2016; Kato et al., 2012). Disruption of these elements abolished their capacity to organize or enter the liquid-liquid separations (Altmeyer et al., 2015). Other molecular elements determining the ability of proteins to enter and exit the nuclear bodies are still elusive. The newly identified Fall protein may serve as a useful model for elucidating the general physiological function of the speckle formation or deformation behavior. Furthermore, Fall is only expressed and functions in body muscles, providing a nice tool to study BTB-ZFs proteins properties and functions in a tissue-specific context.

2.2.4.3 Fall: a homolog of ZBTB42 in fly?

Although BTB-ZF proteins have been found in many organisms, it is hard to determine their evolution path since they normally consist of a long low-complexity region that is very flexible despite of the conserved BTB domain. Interestingly, *ZBTB42*, a BTB zinc finger protein in vertebrates, has been shown to be expressed in skeletal muscles (Devaney et al., 2011). Point mutation in *ZBTB42* leads to lethal congenital contracture syndrome, which is characterized by multiple contractures displayed at distal joints. Moreover, knockdown of *ZBTB42* in zebrafish result in abnormal skeletal muscle development and sarcomere disorders (Patel et al., 2014). Although the underlying mechanism is still elusive, it may employ a similar regulatory path with *fall*. Together, these data suggest that BTB zinc finger proteins (*ZBTB42* and *fall*) represent a new group of regulators of skeletal muscle development. Upon mutation, misregulated downstream targets could induce skeletal muscle hypercontraction. Thus, the mechanistic investigation of Fall may largely extend our understanding of human pathology caused by mutations in *ZBTB42*.

2.3 Outlook

Some questions remain to fully understand the role of Fall during IFM development. First, how Fall does regulate its targets? It may bind to DNA directly and act as a classic transcription factor or work together with chromatin remodelers such as CTCF to modulate chromatin state or architecture. Furthermore, it is unclear that if the localization of Fall to nuclear bodies is coupled to its physiological function. To answer these questions, it would be interesting to perform IP-mass spectrometry to identify its binding partners using Fall-GFP and Fall-2HA as bait. Additionally, mapping the upstream regulatory sequences of Fall targets may provide information about how Fall regulates its targets. Since ChIP-seq technology is still challenging to apply to specific tissue, the novel DNA adenine methyltransferase identification (DamID) method would be a potential alternative (Aughey and Southall, 2016).

Second, what are the upstream regulators that determine the spatial-temporal expression of *fall* and restrict its function specifically to sarcomere maturation stage? *fall* starts to be expressed after myofibrillogenesis initiation (30h APF) and functions largely during sarcomere maturation.

Discussion and Outlook

Third, although Fall is expressed in many body muscles during development, the phenotype was only observed in IFM. How its function is restricted to IFMs is still unknown. One possibility is a IFM-specific interaction partner which is missing in jump muscle and leg muscles may exist. Another possibility is that the targets regulated by Fall are specifically important for sarcomere growth and maturation in IFMs. More detailed analysis of direct targets and interaction partners may help to answer this question.

Finally, it would be also interesting to identify its orthologue in vertebrate if it exists. And thus, the results that we obtained from *Drosophila* may also provide information for understanding sarcomerogenesis in vertebrate. ZBTB42 might be a good candidate.

References:

1. Abmayr, S.M., and Pavlath, G.K. (2012). Myoblast fusion: lessons from flies and mice. *Development* *139*, 641–656.
2. Altmeyer, M., Neelsen, K.J., Teloni, F., Pozdnyakova, I., Pellegrino, S., Grøfte, M., Rask, M.-B.D., Streicher, W., Jungmichel, S., Nielsen, M.L., et al. (2015). Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). *Nat Commun* *6*, 8088.
3. Atreya, K.B., and Fernandes, J.J. (2008). Founder cells regulate fiber number but not fiber formation during adult myogenesis in *Drosophila*. *Dev. Biol.* *321*, 123–140.
4. Aughey, G.N., and Southall, T.D. (2016). Dam it's good! DamID profiling of protein-DNA interactions. *Wiley Interdiscip Rev Dev Biol* *5*, 25–37.
5. Barrangou, R. (2013). CRISPR-Cas systems and RNA-guided interference. *Wiley Interdiscip Rev RNA*.
6. Bassett, A.R., Tibbit, C., Ponting, C.P., and Liu, J.-L. (2013). Highly Efficient Targeted Mutagenesis of *Drosophila* with the CRISPR/Cas9 System. *Cell Reports* *4*, 220–228.
7. Bate, M. (1990). The embryonic development of larval muscles in *Drosophila*. *Development* *110*, 791–804.
8. Bate, M., Rushton, E., and Currie, D.A. (1991). Cells with persistent twist expression are the embryonic precursors of adult muscles in *Drosophila*. *Development* *113*, 79–89.
9. Bergeron-Sandoval, L.-P., Safaee, N., and Michnick, S.W. (2016). Mechanisms and Consequences of Macromolecular Phase Separation. *Cell* *165*, 1067–1079.
10. Bernard, F., Dutriaux, A., Silber, J., and Lalouette, A. (2006). Notch pathway repression by vestigial is required to promote indirect flight muscle differentiation in *Drosophila melanogaster*. *Dev. Biol.* *295*, 164–177.
11. Bernard, F., Lalouette, A., Gullaud, M., Jeantet, A.Y., Cossard, R., Zider, A., Ferveur, J.F., and Silber, J. (2003). Control of apterous by vestigial drives indirect flight muscle development in *Drosophila*. *Dev. Biol.* *260*, 391–403.
12. Bernard, F., Kasherov, P., Grenetier, S., Dutriaux, A., Zider, A., Silber, J., and Lalouette, A. (2009). Developmental Biology. *Dev. Biol.* *332*, 258–272.
13. Black, B.L., and Olson, E.N. (1998). Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. *Annu. Rev. Cell Dev. Biol.* *14*, 167–196.
14. Boateng, S.Y., and Goldspink, P.H. (2007). Assembly and maintenance of the sarcomere night and day. *Cardiovascular Research* *77*, 667–675.
15. Bour, B.A., O'Brien, M.A., Lockwood, W.L., Goldstein, E.S., Bodmer, R., Taghert, P.H., Abmayr, S.M., and Nguyen, H.T. (1995). *Drosophila* MEF2, a transcription factor that is essential for myogenesis. *Genes & Development* *9*, 730–741.

References

16. Bryantsev, A.L., Baker, P.W., Lovato, T.L., Jaramillo, M.S., and Cripps, R.M. (2012a). Differential requirements for Myocyte Enhancer Factor-2 during adult myogenesis in *Drosophila*. *Dev. Biol.* *361*, 191–207.
17. Bryantsev, A.L., Duong, S., Brunetti, T.M., Chechenova, M.B., Lovato, T.L., Nelson, C., Shaw, E., Uhl, J.D., Gebelein, B., and Cripps, R.M. (2012b). Extradenticle and homothorax control adult muscle fiber identity in *Drosophila*. *Dev. Cell* *23*, 664–673.
18. Buckingham, M. (2001). Skeletal muscle formation in vertebrates. *Curr. Opin. Genet. Dev.* *11*, 440–448.
19. Buckingham, M. (2006). Myogenic progenitor cells and skeletal myogenesis in vertebrates. *Curr. Opin. Genet. Dev.* *16*, 525–532.
20. Campbell, K.B., and Chandra, M. (2006). Functions of stretch activation in heart muscle. *J. Gen. Physiol.* *127*, 89–94.
21. Carmena, A., Murugasu-Oei, B., Menon, D., Jiménez, F., and Chia, W. (1998). Inscuteable and numb mediate asymmetric muscle progenitor cell divisions during *Drosophila* myogenesis. *Genes & Development* *12*, 304–315.
22. Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* *39*, 82.
23. Clark, K.A., McElhinny, A.S., Beckerle, M.C., and Gregorio, C.C. (2002). Striated muscle cytoarchitecture: an intricate web of form and function. *Annu. Rev. Cell Dev. Biol.* *18*, 637–706.
24. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* *339*, 819–823.
25. Craig, R., and Lehman, W. (2001). Crossbridge and tropomyosin positions observed in native, interacting thick and thin filaments. *J. Mol. Biol.* *311*, 1027–1036.
26. Cripps, R.M., Becker, K.D., Mardahl, M., Kronert, W.A., Hodges, D., and Bernstein, S.I. (1994). Transformation of *Drosophila melanogaster* with the wild-type myosin heavy-chain gene: rescue of mutant phenotypes and analysis of defects caused by overexpression. *J. Cell Biol.* *126*, 689–699.
27. Cripps, R.M., Black, B.L., Zhao, B., Lien, C.L., Schulz, R.A., and Olson, E.N. (1998). The myogenic regulatory gene Mef2 is a direct target for transcriptional activation by Twist during *Drosophila* myogenesis. *Genes & Development* *12*, 422–434.
28. De Deyne, P.G. (2000). Formation of sarcomeres in developing myotubes: role of mechanical stretch and contractile activation. *AJP: Cell Physiology* *279*, C1801–C1811.
29. de Joussineau, C., Bataillé, L., Jagla, T., and Jagla, K. (2012). Diversification of muscle types in *Drosophila*: upstream and downstream of identity genes. *Curr. Top. Dev. Biol.* *98*, 277–301.
30. De Oliveira Vilaca, L.M. (2013). Mathematical modeling and numerical simulation of excitation-contraction phenomena in the heart.

References

31. Devaney, S.A., Mate, S.E., Devaney, J.M., and Hoffman, E.P. (2011). Characterization of the ZBTB42 gene in humans and mice. *Hum. Genet.* 129, 433–441.
32. Dhordain, P., Albagli, O., Lin, R.J., Ansieau, S., Quief, S., Leutz, A., Kerckaert, J.P., Evans, R.M., and Leprince, D. (1997). Corepressor SMRT binds the BTB/POZ repressing domain of the LAZ3/BCL6 oncoprotein. *Proceedings of the National Academy of Sciences* 94, 10762–10767.
33. Dicarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.*
34. Dietzl, G., Chen, D., Schnorrer, F., Su, K.-C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblaue, S., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448, 151–156.
35. Dobi, K.C., Schulman, V.K., and Baylies, M.K. (2015). Specification of the somatic musculature in *Drosophila*. *Wiley Interdiscip Rev Dev Biol* 4, 357–375.
36. Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346, 1258096.
37. Downes, M., Ordentlich, P., Kao, H.Y., Alvarez, J.G., and Evans, R.M. (2000). Identification of a nuclear domain with deacetylase activity. *Proc. Natl. Acad. Sci. U.S.A.* 97, 10330–10335.
38. Dutta, D., Anant, S., Ruiz-Gomez, M., Bate, M., and VijayRaghavan, K. (2004). Founder myoblasts and fibre number during adult myogenesis in *Drosophila*. *Development* 131, 3761–3772.
39. Ehler, E., and Gautel, M. (2008). The Sarcomere and Sarcomerogenesis. In *Link.Springer.com*, (New York, NY: Springer New York), pp. 1–14.
40. Ejsmont, R.K., Sarov, M., Winkler, S., Lipinski, K.A., and Tomancak, P. (2009). A toolkit for high-throughput, cross-species gene engineering in *Drosophila*. *Nat. Methods* 6, 435–437.
41. Farman, G.P., Miller, M.S., Reedy, M.C., Soto-Adames, F.N., Vigoreaux, J.O., Maughan, D.W., and Irving, T.C. (2009). *Journal of Structural Biology*. *Journal of Structural Biology* 168, 240–249.
42. Fei Xie, L.Y.J.C.C.A.I.B.J.W.M.O.M.Y.W.K. (2014). Seamless gene correction of β -thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Research* 24, 1526.
43. Feng, Q., Lakey, A., Agianian, B., and Hutchings, A. (2003). Troponin C in different insect muscle types: identification of two isoforms in *Lethocerus*, *Drosophila* and *Anopheles* that are specific to asynchronous flight muscle in *Biochemical*
44. Fernandes, J.J., Celniker, S.E., and VijayRaghavan, K. (1996). Development of the indirect flight muscle attachment sites in *Drosophila*: role of the PS integrins and the stripe gene. *Dev. Biol.* 176, 166–184.
45. Figeac, N., Jagla, T., Aradhya, R., Da Ponte, J.P., and Jagla, K. (2010). *Drosophila* adult muscle precursors form a network of interconnected cells and are specified by the rhomboid-triggered EGF pathway. *Development* 137, 1965–1973.

References

46. Geeves, M.A., and Holmes, K.C. (1999). Structural mechanism of muscle contraction. *Annu. Rev. Biochem.* *68*, 687–728.
47. Ghazi, A., and VijayRaghavan, K. (2003). Muscle development in *Drosophila*.
48. Gildor, B., Schejter, E.D., and Shilo, B.Z. (2012). Bidirectional Notch activation represses fusion competence in swarming adult *Drosophila* myoblasts. *Development* *139*, 4040–4050.
49. Gomes, A.V., Potter, J.D., and Szczesna-Cordary, D. (2002). The role of troponins in muscle contraction. *IUBMB Life* *54*, 323–333.
50. Gordon, A.M., Homsher, E., and Regnier, M. (2000). Regulation of Contraction in Striated Muscle. *Physiol. Rev.* *80*, 853–924.
51. Gómez-Díaz, E., and Corces, V.G. (2014). Architectural proteins: regulators of 3D genome organization in cell fate. *Trends in Cell Biology* *24*, 703–711.
52. Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., and O'Connor-Giles, K.M. (2013). Genome Engineering of *Drosophila* with the CRISPR RNA-Guided Cas9 Nuclease. *Genetics* *194*, 1029–1035.
53. Gratz, S.J., Ukken, F.P., Rubinstein, C.D., Thiede, G., Donohue, L.K., Cummings, A.M., and O'Connor-Giles, K.M. (2014). Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics* *196*, 961–971.
54. Gurnett, C.A., Desruisseau, D.M., McCall, K., Choi, R., Meyer, Z.I., Talerico, M., Miller, S.E., Ju, J.S., Pestronk, A., Connolly, A.M., et al. (2010). Myosin binding protein C1: a novel gene for autosomal dominant distal arthrogryposis type 1. *Hum. Mol. Genet.* *19*, 1165–1173.
55. Hanson, J., and Huxley, H.E. (1953). Structural basis of the cross-striations in muscle. *Nature* *172*, 530–532.
56. Herranz, R., Díaz-Castillo, C., Nguyen, T.P., Lovato, T.L., Cripps, R.M., and Marco, R. (2004). Expression patterns of the whole troponin C gene repertoire during *Drosophila* development. *Gene Expr. Patterns* *4*, 183–190.
57. Holtzer, H., Hijikata, T., Lin, Z.X., Zhang, Z.Q., Holtzer, S., Protasi, F., Franzini-Armstrong, C., and Sweeney, H.L. (1997). Independent assembly of 1.6 microns long bipolar MHC filaments and I-Z-I bodies. *Cell Struct. Funct.* *22*, 83–93.
58. Hooper, S.L., and Thuma, J.B. (2005). Invertebrate muscles: muscle specific genes and proteins. *Physiol. Rev.* *85*, 1001–1060.
59. Hooper, S.L., Hobbs, K.H., and Thuma, J.B. (2008). Invertebrate muscles: Thin and thick filament structure; molecular basis of contraction and its regulation, catch and asynchronous muscle. *Prog. Neurobiol.* *86*, 72–127.
60. Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* *157*, 1262–1278.
61. Huxley, A.F., and Niedergerke, R. (1954). Structural changes in muscle during contraction; interference microscopy of living muscle fibres. *Nature* *173*, 971–973.
62. Huxley, H.E. (2004). Fifty years of muscle and the sliding filament hypothesis. *The FEBS Journal* *271*, 1403–1415.

References

63. Huynh, K.D., and Bardwell, V.J. (1998). The BCL-6 POZ domain and other POZ domains interact with the co-repressors N-CoR and SMRT. *Oncogene* *17*, 2473–2484.
64. Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.-R.J., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* *31*, 227–229.
65. Hyman, A.A., Weber, C.A., and Jülicher, F. (2014). Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* *30*, 39–58.
66. Josephson, R.K., Malamud, J.G., and Stokes, D.R. (2000). Asynchronous muscle: a primer. *J. Exp. Biol.* *203*, 2713–2722.
67. Joung, J.K., and Sander, J.D. (2012). TALENs: a widely applicable technology for targeted genome editing. *Nature Reviews Molecular Cell Biology* *14*, 49–55.
68. Kampourakis, T., Sun, Y.-B., and Irving, M. (2016). Myosin light chain phosphorylation enhances contraction of heart muscle via structural changes in both thick and thin filaments. *Proc. Natl. Acad. Sci. U.S.A.* *113*, E3039–E3047.
69. Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., et al. (2012). Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. *Cell* *149*, 753–767.
70. Katsuyama, T., Akmammedov, A., Seimiya, M., Hess, S.C., Sievers, C., and Paro, R. (2013). An efficient strategy for TALEN-mediated genome engineering in *Drosophila*. *Nucleic Acids Res.* *41*, e163–e163.
71. Kaya-Çopur, A., and Schnorrer, F. (2016). A Guide to Genome-Wide In Vivo RNAi Applications in *Drosophila*. In *Methods in Molecular Biology*, (New York, NY: Springer New York), pp. 117–143.
72. Kim, J.H., Jin, P., Duan, R., and Chen, E.H. (2015a). Mechanisms of myoblast fusion during muscle development. *Curr. Opin. Genet. Dev.* *32*, 162–170.
73. Kim, J.H., Ren, Y., Ng, W.P., Li, S., Son, S., Kee, Y.-S., Zhang, S., Zhang, G., Fletcher, D.A., Robinson, D.N., et al. (2015b). Mechanical tension drives cell membrane fusion. *Dev. Cell* *32*, 561–573.
74. Kimber, E., Tajsharghi, H., Kroksmark, A.-K., Oldfors, A., and Tulinius, M. (2006). A mutation in the fast skeletal muscle troponin I gene causes myopathy and distal arthrogryposis. *Neurology* *67*, 597–601.
75. Kimber, E. (2015). AMC: amyoplasia and distal arthrogryposis. *J Child Orthop* *9*, 427–432.
76. Koken, M.H.M., Reid, A., Quignon, F., Chelbi-Alix, M.K., Davies, J.M., Kabarowski, J.H.S., Zhu, J., Dong, S., Chen, S.J., Chen, Z., et al. (1997). Leukemia-associated retinoic acid receptor α fusion partners, PML and PLZF, heterodimerize and colocalize to nuclear bodies. *Proceedings of the National Academy of Sciences* *94*, 10255–10260.
77. Kontogianni-Konstantopoulos, A., Ackermann, M.A., Bowman, A.L., Yap, S.V., and Bloch, R.J. (2009). Muscle giants: molecular scaffolds in sarcomerogenesis. *Physiol. Rev.* *89*, 1217–1267.

References

78. Kržič, U., Rybin, V., Leonard, K.R., Linke, W.A., and Bullard, B. (2010). Regulation of Oscillatory Contraction in Insect Flight Muscle by Troponin. *J. Mol. Biol.* *397*, 110–118.
79. Lemke, S.B., and Schnorrer, F. (2016). Mechanical forces during muscle development. *Mech. Dev.*
80. Linari, M., Reedy, M.K., Reedy, M.C., Lombardi, V., and Piazzesi, G. (2004). Ca-activation and stretch-activation in insect flight muscle. *Biophys. J.* *87*, 1101–1111.
81. Liu, J., Li, C., Yu, Z., Huang, P., Wu, H., Wei, C., Zhu, N., Shen, Y., Chen, Y., Zhang, B., et al. (2012). Efficient and Specific Modifications of the *Drosophila* Genome by Means of an Easy TALEN Strategy. *Journal of Genetics and Genomics* *39*, 209–215.
82. Mao, Y.S., Zhang, B., and Spector, D.L. (2011). Biogenesis and function of nuclear bodies. *Trends Genet.* *27*, 295–306.
83. Markus, B., Narkis, G., Landau, D., Birk, R.Z., Cohen, I., and Birk, O.S. (2012). Autosomal recessive lethal congenital contractural syndrome type 4 (LCCS4) caused by a mutation in MYBPC1. *Hum. Mutat.* *33*, 1435–1438.
84. Mendell, J.R., and Rodino-Klapac, L.R. (2016). Duchenne muscular dystrophy: CRISPR/Cas9 treatment. *Cell Res.* *26*, 513–514.
85. Mitrea, D.M., and Kriwacki, R.W. (2016). Phase separation in biology; functional organization of a higher order. *Cell Commun. Signal* *14*, 1.
86. Montana, E.S., and Littleton, J.T. (2004). Characterization of a hypercontraction-induced myopathy in *Drosophila* caused by mutations in Mhc. *J. Cell Biol.* *164*, 1045–1054.
87. Moore, J.R. (2006). Stretch Activation. In Link.Springer.com, (Boston, MA: Springer US), pp. 44–60.
88. Mukherjee, P., Gildor, B., Shilo, B.Z., VijayRaghavan, K., and Schejter, E.D. (2011). The actin nucleator WASp is required for myoblast fusion during adult *Drosophila* myogenesis. *Development* *138*, 2347–2357.
89. Nagarkar-Jaiswal, S., Lee, P.-T., Campbell, M.E., Chen, K., Anguiano-Zarate, S., Gutierrez, M.C., Busby, T., Lin, W.-W., He, Y., Schulze, K.L., et al. (2015). A library of MiMICs allows tagging of genes and reversible, spatial and temporal knockdown of proteins in *Drosophila*. *Elife* *4*, e05338.
90. Nongthomba, U., Ansari, M., Thimmaiya, D., Stark, M., and Sparrow, J. (2007). Aberrant Splicing of an Alternative Exon in the *Drosophila* Troponin-T Gene Affects Flight Muscle Development. *Genetics* *177*, 295–306.
91. Nongthomba, U., Clark, S., Cummins, M., Ansari, M., Stark, M., and Sparrow, J.C. (2004). Troponin I is required for myofibrillogenesis and sarcomere formation in *Drosophila* flight muscle. *J. Cell. Sci.* *117*, 1795–1805.
92. Nongthomba, U., Cummins, M., Clark, S., Vigoreaux, J.O., and Sparrow, J.C. (2003). Suppression of muscle hypercontraction by mutations in the myosin heavy chain gene of *Drosophila melanogaster*. *Genetics* *164*, 209–222.
93. Oas, S.T., Bryantsev, A.L., and Cripps, R.M. (2014). Arrest is a regulator of fiber-specific alternative splicing in the indirect flight muscles of *Drosophila*. *J. Cell Biol.* jcb.201405058.

References

94. Ohtsuki, I., and Morimoto, S. (2008). Troponin: regulatory function and disorders. *Biochemical and Biophysical Research Communications* 369, 62–73.
95. Ojima, K., Lin, Z.X., Zhang, Z.Q., Hijikata, T., Holtzer, S., Labeit, S., Sweeney, H.L., and Holtzer, H. (1999). Initiation and maturation of I-Z-I bodies in the growth tips of transfected myotubes. *J. Cell. Sci.* 112 (Pt 22), 4101–4112.
96. Orfanos, Z., Leonard, K., Elliott, C., Katzemich, A., Bullard, B., and Sparrow, J. (2015). Sallimus and the Dynamics of Sarcomere Assembly in *Drosophila* Flight Muscles. *J. Mol. Biol.* 427, 2151–2158.
97. Patel, N., Smith, L.L., Fageih, E., Mohamed, J., Gupta, V.A., and Alkuraya, F.S. (2014). ZBTB42 mutation defines a novel lethal congenital contracture syndrome (LCCS6). *Hum. Mol. Genet.*
98. Perz-Edwards, R.J., Irving, T.C., Baumann, B.A.J., Gore, D., Hutchinson, D.C., Kržič, U., Porter, R.L., Ward, A.B., and Reedy, M.K. (2011). X-ray diffraction evidence for myosin-troponin connections and tropomyosin movement during stretch activation of insect flight muscle. *Proceedings of the National Academy of Sciences* 108, 120–125.
99. Pringle, J.W. (1949). The excitation and contraction of the flight muscles of insects. *The Journal of Physiology* 108, 226–232.
100. Ranganayakulu, G., Zhao, B., Dokidis, A., Molkentin, J.D., Olson, E.N., and Schulz, R.A. (1995). A series of mutations in the D-MEF2 transcription factor reveal multiple functions in larval and adult myogenesis in *Drosophila*. *Dev. Biol.* 171, 169–181.
101. Reedy, M.C., and Beall, C. (1993). Ultrastructure of developing flight muscle in *Drosophila*. I. Assembly of myofibrils. *Dev. Biol.* 160, 443–465.
102. Reid, A., Gould, A., Brand, N., Cook, M., Strutt, P., Li, J., Licht, J., Waxman, S., Krumlauf, R., and Zelent, A. (1995). Leukemia translocation gene, PLZF, is expressed with a speckled nuclear pattern in early hematopoietic progenitors. *Blood* 86, 4544–4552.
103. Rivlin, P.K., Schneiderman, A.M., and Booker, R. (2000). Imaginal Pioneers Prefigure the Formation of Adult Thoracic Muscles in *Drosophila melanogaster*. *Dev. Biol.* 222, 450–459.
104. Robinson, P., Lipscomb, S., Preston, L.C., Altin, E., Watkins, H., Ashley, C.C., and Redwood, C.S. (2007). Mutations in fast skeletal troponin I, troponin T, and beta-tropomyosin that cause distal arthrogryposis all increase contractile function. *The FASEB Journal* 21, 896–905.
105. Roy, S., and VijayRaghavan, K. (1999). Muscle pattern diversification in *Drosophila*: the story of imaginal myogenesis. *Bioessays* 21, 486–498.
106. Rubin, G.M. (2001). The draft sequences: Comparing species. *Nature* 409, 820–821.
107. Ruiz Gomez, M., and Bate, M. (1997). Segregation of myogenic lineages in *Drosophila* requires numb. *Development* 124, 4857–4866.
108. Sanders, K.M. (2008). Regulation of smooth muscle excitation and contraction. *Neurogastroenterol. Motil.* 20 Suppl 1, 39–53.

References

109. Sanger, J.W., Kang, S., Siebrands, C.C., Freeman, N., Du, A., Wang, J., Stout, A.L., and Sanger, J.M. (2005). How to build a myofibril. *J. Muscle Res. Cell. Motil.* **26**, 343–354.
110. Sanger, J.W., Wang, J., Fan, Y., White, J., Mi-Mi, L., Dube, D.K., Sanger, J.M., and Pruyne, D. (2017). Assembly and Maintenance of Myofibrils in Striated Muscle. *Handb Exp Pharmacol* **235**, 39–75.
111. Sarov, M., Barz, C., Jambor, H., Hein, M.Y., Schmied, C., Suchold, D., Stender, B., Janosch, S., K J, V.V., Krishnan, R.T., et al. (2016). A genome-wide resource for the analysis of protein localisation in *Drosophila*. *Elife* **5**, e12068.
112. Schejter, E.D. (2017). Seminars in Cell & Developmental Biology. *Seminars in Cell & Developmental Biology* 1–9.
113. Schnorrer, F., and Dickson, B.J. (2004). Muscle building; mechanisms of myotube guidance and attachment site selection. *Dev. Cell* **7**, 9–20.
114. Schnorrer, F., Schönbauer, C., Langer, C.C.H., Dietzl, G., Novatchkova, M., Schernhuber, K., Fellner, M., Azaryan, A., Radolf, M., Stark, A., et al. (2010). Systematic genetic analysis of muscle morphogenesis and function in *Drosophila*. *Nature* **464**, 287–291.
115. Schoborg, T., Rickels, R., Barrios, J., and Labrador, M. (2013). Chromatin insulator bodies are nuclear structures that form in response to osmotic stress and cell death. *J. Cell Biol.* **202**, 261–276.
116. Schönbauer, C., Distler, J., Jährling, N., Radolf, M., Dodt, H.-U., Frasch, M., and Schnorrer, F. (2011). Spalt mediates an evolutionarily conserved switch to fibrillar muscle fate in insects. *Nature* **479**, 406–409.
117. Schönbauer, C., Distler, J., Jährling, N., Radolf, M., Dodt, H.-U., Frasch, M., and Schnorrer, F. (2012). Spalt mediates an evolutionarily conserved switch to fibrillar muscle fate in insects. *Nature* **479**, 406–409.
118. Schulman, V.K., Dobi, K.C., and Baylies, M.K. (2015). Morphogenesis of the somatic musculature in *Drosophila melanogaster*. *Wiley Interdiscip Rev Dev Biol* **4**, 313–334.
119. Shav-Tal, Y., Blechman, J., Darzacq, X., Montagna, C., DYE, B.T., PATTON, J.G., Singer, R.H., and Zipori, D. (2005). Dynamic sorting of nuclear components into distinct nucleolar caps during transcriptional inhibition. *Mol. Biol. Cell* **16**, 2395–2413.
120. Shwartz, A., Dhanyasi, N., Schejter, E.D., and Shilo, B.-Z. (2016). The *Drosophila* formin Fhos is a primary mediator of sarcomeric thin-filament array assembly. *Elife* **5**.
121. Siggs, O.M., and Beutler, B. (2012). Cell Cycle: Extra Views. *Cell Cycle*.
122. Simpson, D.G., Decker, M.L., Clark, W.A., and Decker, R.S. (1993). Contractile activity and cell-cell contact regulate myofibrillar organization in cultured cardiac myocytes. *J. Cell Biol.* **123**, 323–336.
123. Soler, C., Han, J., and Taylor, M.V. (2012). The conserved transcription factor Mef2 has multiple roles in adult *Drosophila* musculature formation. *Development* **139**, 1270–1275.

References

124. Soler, C., and Taylor, M.V. (2009). The Him gene inhibits the development of *Drosophila* flight muscles during metamorphosis. *Mod* 126, 595–603.
125. Spletter, M.L., and Schnorrer, F. (2014). Transcriptional regulation and alternative splicing cooperate in muscle fiber-type specification in flies and mammals. *Exp. Cell Res.* 321, 90–98.
126. Spletter, M.L., Barz, C., Yeroslaviz, A., Schönbauer, C., Ferreira, I.R.S., Sarov, M., Gerlach, D., Stark, A., Habermann, B.H., and Schnorrer, F. (2015). The RNA-binding protein Arrest (Bruno) regulates alternative splicing to enable myofibril maturation in *Drosophila* flight muscle. *EMBO Rep.* 16, 178–191.
127. Spudich, J.A. (2001). The myosin swinging cross-bridge model. *Nature Reviews Molecular Cell Biology* 2, 387–392.
128. Squire, J.M. (2016). Muscle contraction: Sliding filament history, sarcomere dynamics and the two Huxleys. *Gesp* 2016.
129. Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507, 62–67.
130. Sudarsan, V., Anant, S., Guptan, P., VijayRaghavan, K., and Skaer, H. (2001). Myoblast diversification and ectodermal signaling in *Drosophila*. *Dev. Cell* 1, 829–839.
131. Sung, S.S., Brassington, A.-M.E., Grannatt, K., Rutherford, A., Whitby, F.G., Krakowiak, P.A., Jorde, L.B., Carey, J.C., and Bamshad, M. (2003). Mutations in Genes Encoding Fast-Twitch Contractile Proteins Cause Distal Arthrogryposis Syndromes. *The American Journal of Human Genetics* 72, 681–690.
132. Sung, Y.H., Baek, I.-J., Kim, D.H., Jeon, J., Lee, J., Lee, K., Jeong, D., Kim, J.-S., and Lee, H.-W. (2013). Knockout mice created by TALEN-mediated gene targeting. *Nat. Biotechnol.* 31, 23–24.
133. Szczesna, D., Zhao, J., Jones, M., Zhi, G., Stull, J., and Potter, J.D. (2002). Phosphorylation of the regulatory light chains of myosin affects Ca²⁺ sensitivity of skeletal muscle contraction. *J. Appl. Physiol.* 92, 1661–1670.
134. Tajsharghi, H., Kimber, E., Holmgren, D., Tulinius, M., and Oldfors, A. (2007). Distal arthrogryposis and muscle weakness associated with a beta-tropomyosin mutation. *Neurology* 68, 772–775.
135. Takashima, S. (2009). Phosphorylation of myosin regulatory light chain by myosin light chain kinase, and muscle contraction. *Circ. J.* 73, 208–213.
136. Taylor, M.V. (2006). Comparison of Muscle Development in *Drosophila* and Vertebrates. In *Muscle Development in Drosophila*, (New York, NY: Springer New York), pp. 169–203.
137. Themeli, M., Rivière, I., and Sadelain, M. (2015). New Cell Sources for T Cell Engineering and Adoptive Immunotherapy. *Cell Stem Cell* 16, 357–366.
138. Tixier, V., Bataillé, L., and Jagla, K. (2010). Diversification of muscle types: recent insights from *Drosophila*. *Exp. Cell Res.* 316, 3019–3027.
139. Toydemir, R.M., Chen, H., Proud, V.K., Martin, R., van Bokhoven, H., Hamel, B.C.J., Tuerlings, J.H., Stratakis, C.A., Jorde, L.B., and Bamshad, M.J. (2006a).

References

- Trismus-pseudocamptodactyly syndrome is caused by recurrent mutation of MYH8. *Am. J. Med. Genet. 140A*, 2387–2393.
140. Toydemir, R.M., Rutherford, A., Whitby, F.G., Jorde, L.B., Carey, J.C., and Bamshad, M.J. (2006b). Mutations in embryonic myosin heavy chain (MYH3) cause Freeman-Sheldon syndrome and Sheldon-Hall syndrome. *Nat. Genet.* 38, 561–565.
 141. Ugur, B., Chen, K., and Bellen, H.J. (2016). Drosophila tools and assays for the study of human diseases. *Disease Models and Mechanisms* 9, 235–244.
 142. Venken, K.J.T., Schulze, K.L., Haelterman, N.A., Pan, H., He, Y., Evans-Holm, M., Carlson, J.W., Levis, R.W., Spradling, A.C., Hoskins, R.A., et al. (2011). MiMIC: a highly versatile transposon insertion resource for engineering *Drosophila melanogaster* genes. *Nat. Methods* 8, 737–743.
 143. Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910–918.
 144. Wangler, M.F., Yamamoto, S., and Bellen, H.J. (2015). Fruit Flies in Biomedical Research. *Genetics* 199, 639–653.
 145. Webb, R.C., and Webb, R.C. (2003). Smooth muscle contraction and relaxation. *Adv Physiol Educ* 27, 201–206.
 146. Weitkunat, M., Kaya-Çopur, A., Grill, S.W., and Schnorrer, F. (2014). Tension and force-resistant attachment are essential for myofibrillogenesis in *Drosophila* flight muscle. *Curr. Biol.* 24, 705–716.
 147. Weitkunat, M., Lindauer, M., Bausch, A., and Schnorrer, F. (2017). Mechanical tension and spontaneous muscle twitching precede the formation of cross-striated muscle in vivo. *Development* dev.140723.
 148. Zappia, M.P., and Frolov, M.V. (2016). E2F function in muscle growth is necessary and sufficient for viability in *Drosophila*. *Nat Commun* 7, 10509.
 149. Zhang, X., Ferreira, I.R.S., and Schnorrer, F. (2014a). A simple TALEN-based protocol for efficient genome-editing in *Drosophila*. *Methods* 69, 32–37.
 150. Zhang, X., Koolhaas, W.H., and Schnorrer, F. (2014b). A Versatile Two-Step CRISPR- and RMCE-Based Strategy for Efficient Genome Engineering in *Drosophila*. *G3 (Bethesda)* 4, 2409–2418.
 151. Zhu, L., and Brangwynne, C.P. (2015). Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Current Opinion in Cell Biology* 34, 23–30.
 152. Zu, Y., Tong, X., Wang, Z., Liu, D., Pan, R., Li, Z., Hu, Y., Luo, Z., Huang, P., Wu, Q., et al. (2013). TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat. Methods*.

Appendices

Publication I

Publication II

Manuscript I

Appendices

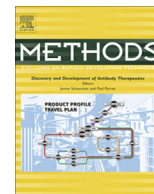
Publication I

A simple TALEN-based protocol for efficient genome-editing in *Drosophila*

Xu Zhang, Irene R.S. Ferreira, Frank Schnorrer

Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

Keywords: *Drosophila* Mutagenesis, Genome-editing, TALEN, T7 endonuclease I



A simple TALEN-based protocol for efficient genome-editing in *Drosophila*



Xu Zhang, Irene R.S. Ferreira, Frank Schnorrer*

Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

ARTICLE INFO

Article history:

Received 6 December 2013

Revised 11 March 2014

Accepted 17 March 2014

Available online 26 March 2014

Keywords:

Drosophila

Mutagenesis

Genome-editing

TALEN

T7 endonuclease I

Method

ABSTRACT

Drosophila is a well-established genetic model organism: thousands of point mutations, deficiencies or transposon insertions are available from stock centres. However, to date, it is still difficult to modify a specific gene locus in a defined manner. A potential solution is the application of transcription activator-like effector nucleases (TALENs), which have been used successfully to mutate genes in various model organisms. TALENs are constructed by fusion of TALE proteins to the endonuclease FokI, resulting in artificial, sequence-specific endonucleases. They induce double strand breaks, which are either repaired by error-prone non-homologous end joining (NHEJ) or homology directed repair (HDR). We developed a simple TALEN-based protocol to mutate any gene of interest in *Drosophila* within approximately 2 months. We inject mRNA coding for two TALEN pairs targeting the same gene into embryos, employ T7 endonuclease I screening of pooled F₁ flies to identify mutations and generate a stable mutant stock in the F₃ generation. We illustrate the efficacy of our strategy by mutating CG11617, a previously uncharacterized putative transcription factor with an unknown function in *Drosophila*. This demonstrates that TALENs are a reliable and efficient strategy to mutate any gene of interest in *Drosophila*.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Drosophila melanogaster is a classical model to investigate gene function in a complex, multi-cellular organism. Starting with the legendary Heidelberg screen for embryonic patterning [1,2] and followed by hundreds of its derivatives, forward genetic screens have been immensely powerful to discover genes and study their functions in *Drosophila* biology (reviewed in [3]). Forward screens per definition hit the genome at random and thus discover new roles for genes in a particular biological context.

Reverse genetics, on the other hand, tests gene function by a targeted approach manipulating one particular gene of interest. After 30 years of forward screens, reverse genetics is currently booming in *Drosophila*, as many large scale systematic approaches such as interaction proteomics [4] or genome-wide RNAi screens [5–7] suggest a function for a given gene without providing classical loss of function mutations. Tools to apply reverse genetics in *Drosophila* have advanced rapidly in the last years, in particular through gene targeting by homologous recombination that allows gene modification at the endogenous gene locus, similar to the ap-

proach in mice [8,9]. However, despite significant progress in the gene targeting protocol enabling smart genetic selection methods [10], it remains relatively tedious to engineer a particular gene of interest to either a genetic null allele or a tagged protein allele. The recent implementation of zinc-finger nucleases in *Drosophila* opened a promising new path for gene engineering [11,12], but zinc-finger design is laborious and has certain sequence restrictions. Mixed-success reports, most likely dependent on the target sequence [13], have prevented the technology from taking off as a routine application for generating mutants in the average fly lab.

Transcription activator-like effector nucleases (TALENs) are derived from naturally occurring transcription factors present in plant pathogens called TALEs, which contain 12–27 repeats, typically composed of 34 amino acids. Each repeat binds to a single base pair of DNA, with the two central variable amino acids determining base specificity [14,15]. In TALENs, the C-terminal transcriptional activation domain has been replaced by the FokI nuclease domain, converting a transcriptional activator to a sequence specific nuclease. As FokI cleaves only as a dimer, TALENs are used in pairs that bind on opposite DNA strands, separated by a central spacer domain. TALEN-induced double strand DNA breaks are either repaired by non-homologous end joining (NHEJ),

* Corresponding author.

E-mail address: schnorrer@biochem.mpg.de (F. Schnorrer).

which often results in small insertions or deletions and is thus valuable to mutate genes, or by homology directed repair (HDR) using the homologous chromosome or a provided homologous DNA sequence, which can be used for gene engineering [16]. This strategy has been applied to a number of different organisms or cell types, including *Arabidopsis*, zebrafish, and human cell lines, to mutate genes or insert DNA sequences into a locus of interest [17–20]. In addition to their straight forward application, TALENs are more flexible than zinc-fingers as they can recognise any

DNA sequence and also appear to have higher target specificity and lower toxicity [19,21,22].

In *Drosophila*, TALENs have mainly been used to mutate genes that result in visible phenotypes such as changes in eye or body colour. Thus, they can easily be scored in the F_1 generation [22–24]. Lately, this technology has been expanded to also mutate genes with unknown phenotypes in *Drosophila* [22] or to insert GFP into desired locations in the fly genome [24], demonstrating the potential of TALEN-mediated genome editing. Here, we report

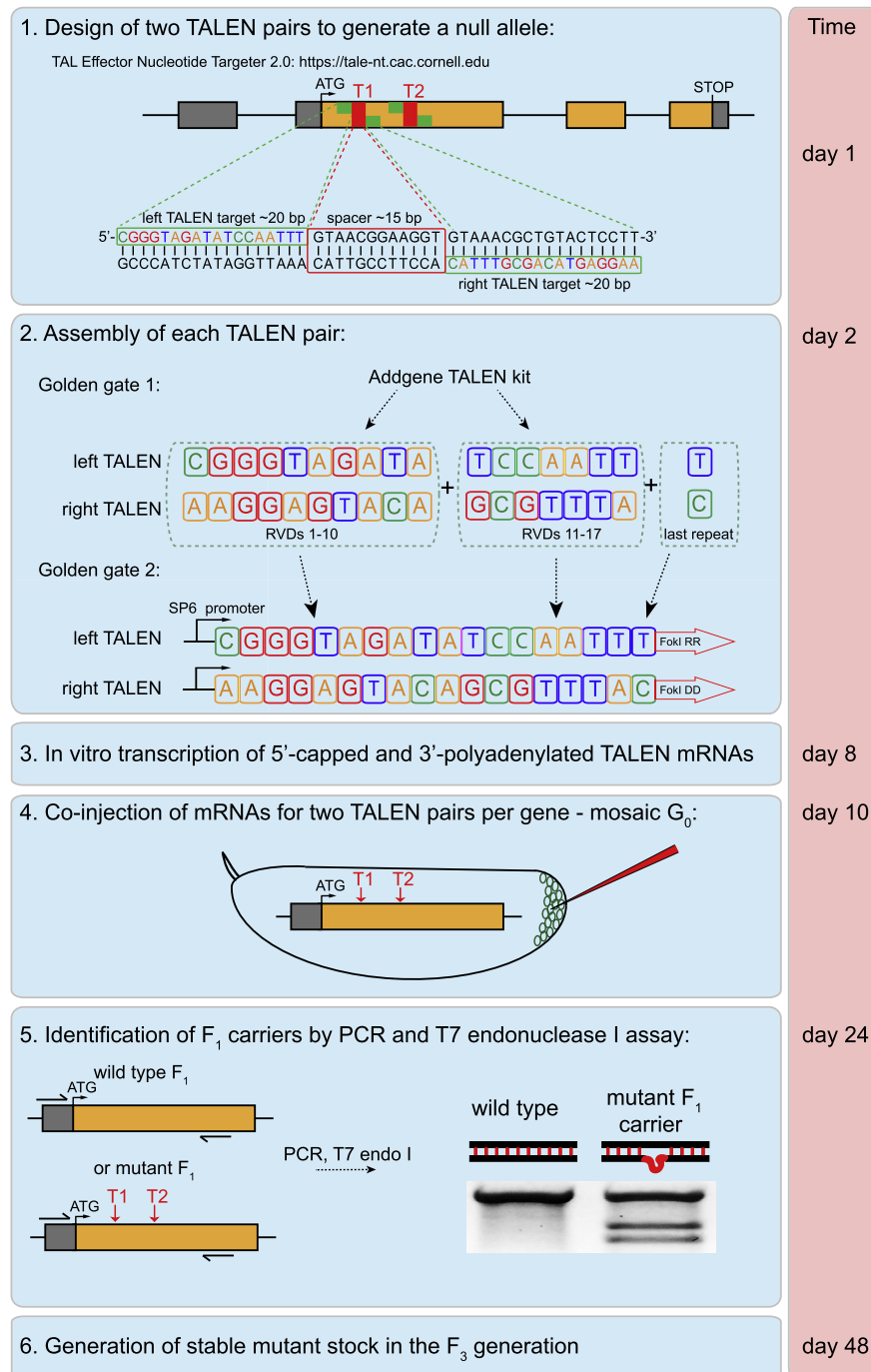


Fig. 1. Overview of TALEN application to generate *Drosophila* mutants. Blue boxes highlight the 6 distinct steps in the procedure to generate a null allele for a given *Drosophila* gene by the application of TALENs. These steps include the design of TALEN pair target sequences within the first coding exon (1); assembly of each TALEN pair using the Golden Gate cloning kit (2); *in vitro* transcription, polyadenylation and capping of TALEN mRNAs (3); co-injection of mRNAs for two TALEN pairs into fly embryos (4); identification of F_1 carriers by T7 endonuclease I screening (5); and generation of stable mutant F_3 stocks (6). The red box estimates the required time for each step.

a simple and effective protocol to apply TALEN-mediated genome engineering in order to mutate any gene of choice in *Drosophila*. Our protocol can easily be performed by a standard *Drosophila* laboratory without the need to purchase special equipment. We generate mosaic G_0 adult flies by injecting *Drosophila* embryos with mRNAs coding for two pairs of TALENs that target next to each other in the same gene region. We apply a T7 endonuclease I screening assay to detect heterozygous mutant carriers amongst the F_1 generation and generate a mutant stock using the F_3 generation. Together, this allows the effective generation of mutations in any *Drosophila* gene within about 2 months.

2. Materials and methods

2.1. Cloning of TALEN pairs

Addgene TALEN kit (Addgene).
pCS2TAL3-DD/RR backbone plasmid (Addgene).
QIAprep Spin Miniprep Kit (Qiagen).
QIAGEN Plasmid Midi Kit (Qiagen).
Bsa I and BsmB I (NEB).
T4 ligase (NEB).
Plasmidsafe nuclease (Epicentre).
Tetracycline 10 μ g/ml (Sigma–Aldrich).
Spectinomycin 50 μ g/ml (Sigma–Aldrich).
Ampicillin 100 μ g/ml (Sigma–Aldrich).

2.2. Plasmid linearisation and in vitro transcription of TALEN pairs

Not I (NEB).
10% SDS (Sigma–Aldrich).
Protease K (Sigma–Aldrich).
Phenol/Chloroform/Isopropanol (v:v:v = 25:24:1).
3 M NaOAc (pH 5.2).
RNase free Eppendorf tubes (1.5 ml) and tips.
mMESSAGE mMACHINE[®] T7 ULTRA Kit (Ambion).
mMESSAGE mMACHINE[®] SP6 Kit (Ambion).

2.3. Embryo injections

Klorix (bleach) to dechorionate the embryos.
Scotch tape dissolved in heptane to glue the dechorionated embryos on a slide for injection.
Silica gel orange (Roth) to dry dechorionated embryos before injection.
10 S Voltalef oil (Lehmann and Voss) to cover the dechorionated embryos during injections.
Glass capillaries (Harvard Apparatus) and needle puller (Sutter Instruments) to pull injection needles.
Eppendorf FemtoJet for injections.

2.4. T7 endonuclease I screening

Fly genomic DNA extraction buffer A (0.1 M Tris, pH 9.0; 0.1 M EDTA; 1% SDS; 5 M KOAc, pH 4.8).
Taq polymerase (NEB).
T7 endonuclease I (NEB).
96-well plates.

3. Application of TALENs to *Drosophila*

Fig. 1 illustrates the different steps in our protocol, starting with design of the TALEN target sequences and ending with the final

mutant *Drosophila* stock. The entire protocol can be completed in about 2 months.

3.1. Design of two TALEN pairs to modify a gene

Identify a region of about 0.2–1 kb in your target gene which you intend to modify, and confirm its sequence in the fly strain that will be used for the engineering experiment (here *white*¹¹¹⁸). To generate a genetic null allele of a given gene, the first coding exon is often a convenient target (Fig. 1–1). Use the TAL Effector Nucleotide Targeter 2.0 (<https://tale-nt.cac.cornell.edu>) program to design suitable TALEN pairs for the target region by using the following recommended parameters: minimum spacer length = 15, maximum spacer length = 16, minimum repeat array length = 18, maximum repeat array length = 20, upstream base = T only, off-target sequence = *D. melanogaster* genome [25]. Rank the outcome for the lowest off-target predictions and choose two pairs of TALENs that do not overlap and thus can create independent mutations at two different positions in the target gene.

3.2. Assembly of TALEN pairs

Assemble four TALENs (two pairs) for each gene region that should be mutated using the Golden Gate kit and its assembly protocol in a standard 20 μ l reaction. All required plasmids are available from Addgene [17] and the detailed protocol can be downloaded from <http://www.addgene.org/static/cms/files/GoldenGateTALAssembly2011.pdf>. In the first assembly step, repeat-variable di-residues (RVDs) 1–10 and 11–20 are assembled (Fig. 1–2). In the second assembly step, both assembled units are combined with the last repeat. In the final step, clone the assembled TALENs into pCS2-TAL3-DD and RR, a convenient vector with an SP6 RNA polymerase promoter and a SV-40 termination site for efficient *in vitro* transcription [19].

3.3. In vitro transcription, capping and poly-adenylation of TALEN pairs

Linearise the TALEN vectors obtained in step 3.2 by standard methods using Not I digestion under RNase free conditions. Using the mMESSAGE mMACHINE T7 Ultra and SP6 kits, assemble a 10 μ l *in vitro* transcription reaction in the following order at room temperature:

Reagent	Volume per reaction (μ l)	Final concentration
2 \times NTP/ARCA (T7 Ultra kit)	5	7.5 mM ATP; 7.5 mM CTP; 7.5 mM UTP; 1.5 mM GTP; 6 mM ARCA
Linearised template	3	~600ng
10 \times reaction buffer (SP6 kit)	1	1 \times
SP6 enzyme mix	1	

This combination of both Ambion kits enables SP6-driven transcription combined with ARCA capping. Incubate at 37 °C overnight.

Add 0.5 μ l TURBO DNase (2 U/ μ l) and incubate at 37 °C for 20 min. Assemble the polyadenylation reaction using the mMESSAGE mMACHINE T7 Ultra kit in the following order:

Reagent	Volume per reaction (μl)	Final concentration
Transcription product	10	~10–15 μg
Nuclease-free water	18	
5 × E-PAP buffer	10	1 ×
25 mM MnCl ₂	5	2.5 mM
10 mM ATP solution	5	1 mM
E-PAP	2	0.08 U μl ⁻¹

Incubate at 37 °C for 45 min, then stop the reaction and precipitate the RNA with 30 μl LiCl (7.5 M). Incubate for 30 min at –20 °C, centrifuge, wash with 70% EtOH, dry and re-suspend the RNA in 10 μl RNase free water. This should produce 10–15 μg ARCA capped polyA mRNA that can be analysed for integrity on a gel and is ready for embryo injections.

3.4. Embryo injection of TALEN pairs

For each target gene locus, inject capped mRNA of two TALEN pairs mixed in water at concentrations of 250 ng/μl for each TALEN mRNA leading to a total concentration of 1 μg/μl. Such high TALEN mRNA concentrations result in higher G₀ lethality, ultimately reducing the following fly work. Inject ~400 *white*¹¹¹⁸ embryos at the posterior pole before germ-cell formation using standard injection procedures. This usually results in about 30–60 viable and fertile G₀ mosaic adults (Fig. 2-1).

3.5. Identification of F₁ carriers by T7 endonuclease I assay

Collect mosaic G₀ males and virgins and cross them singly to balancer flies. This normally is in the range of 30–60 G₀ crosses per gene. As the G₀ flies are mosaic, their F₁ offspring may or may not carry the desired mutation. It is most efficient to pool F₁ flies to identify which carry an inheritable mutation. For each G₀ cross, set up 5 vials with 8 F₁ males per vial and cross them with more than 10 balancer virgins (F₂ cross, Fig. 2-2). After 4 days, recover the 8 mated F₁ males from each pool, fill each pool into single wells of a 96 well plate and homogenise all flies carefully. Extract genomic DNA as published previously [26].

The presence of a mutation in the desired gene within each F₁ pool is determined by T7 endonuclease I assay. Set up a standard 20 μl PCR from each isolated DNA pool amplifying 0.5–1 kb of the gene region including both TALEN target regions (Fig. 1-5):

Reagent	Volume per reaction (μl)	Final concentration
10 × PCR buffer	2	1 ×
50 mM MgCl ₂	0.2	2 mM (total Mg ²⁺)
10 mM dNTPs	0.5	250 μM
10 μM primer mix	0.4	200 nM each primer
Polymerase	0.15	0.0375 U μl ⁻¹
DNA template	1	~300 ng
H ₂ O to final volume of 20 μl	15.75	

Perform the PCR and final denaturation and annealing steps using the following program:

Step	Temperature (°C)	Time
Initial denaturation	95	5 min
Denaturation	95	30 s
Annealing	64	45 s
Extension	68	60 s
<i>Repeat 30 cycles</i>		
Final extension	68	5 min
Final denaturation	95	5 min
Final annealing step 1	95–85	–2 °C/s
Final annealing step 2	85–25	–0.1 °C/s
Hold	4	Indefinitely

PCR products that are a mixture of a mutated allele (base pair insertions or deletions) and a wild-type allele result in a DNA mismatch at the TALEN target site and will be cleaved by T7 endonuclease I using the following reaction mix that is assembled on ice:

Reagent	Volume per reaction (μl)	Final concentration
10 × T7 buffer	2	1 ×
5 U/μl T7 endonuclease I	0.5	0.125 U μl ⁻¹
Annealed PCR product	10	~600 ng
H ₂ O to final volume of 20 μl	7.5	

Incubate the digestion 15–20 min at 37 °C and analyse the products on a 1.5% agarose gel immediately to avoid over-digestion (Figs. 1-5 and 2-2).

Reagent	Volume per reaction (μl)	Final concentration
10 × T7 buffer	2	1 ×
5 U/μl T7 endonuclease I	0.5	0.125 U μl ⁻¹
Annealed PCR product	10	~600 ng
H ₂ O to final volume of 20 μl	7.5	

3.6. Recovery of mutant alleles and generation of stable mutant stock

From each F₁ vial that is positive in the T7 endonuclease I assay, cross 20 single F₂ males to balancer virgins and test the F₃ offspring with the same T7 endonuclease I assay. Upon obtaining a positive T7 result, a mutant stock can be generated from the remaining F₃ males and virgins (Fig. 2-3). Finally, sequence the target gene region of the heterozygous stock to determine the nature of the new mutation.

4. Results and discussion

We applied our protocol to generate null alleles of the putative Zn-finger transcription factor *CG11617* (Fig. 3-1). *CG11617* is located at 21B4 on the second chromosome and mutant alleles have not been described. Following the above protocol, we designed two TALEN pairs targeting the first coding exon of *CG11617* (Fig. 3-2). We injected a mixture of all four TALEN mRNAs into 400 *white*¹¹¹⁸ embryos, resulting in 36 fertile G₀ flies that we crossed with *lfl/CyO*

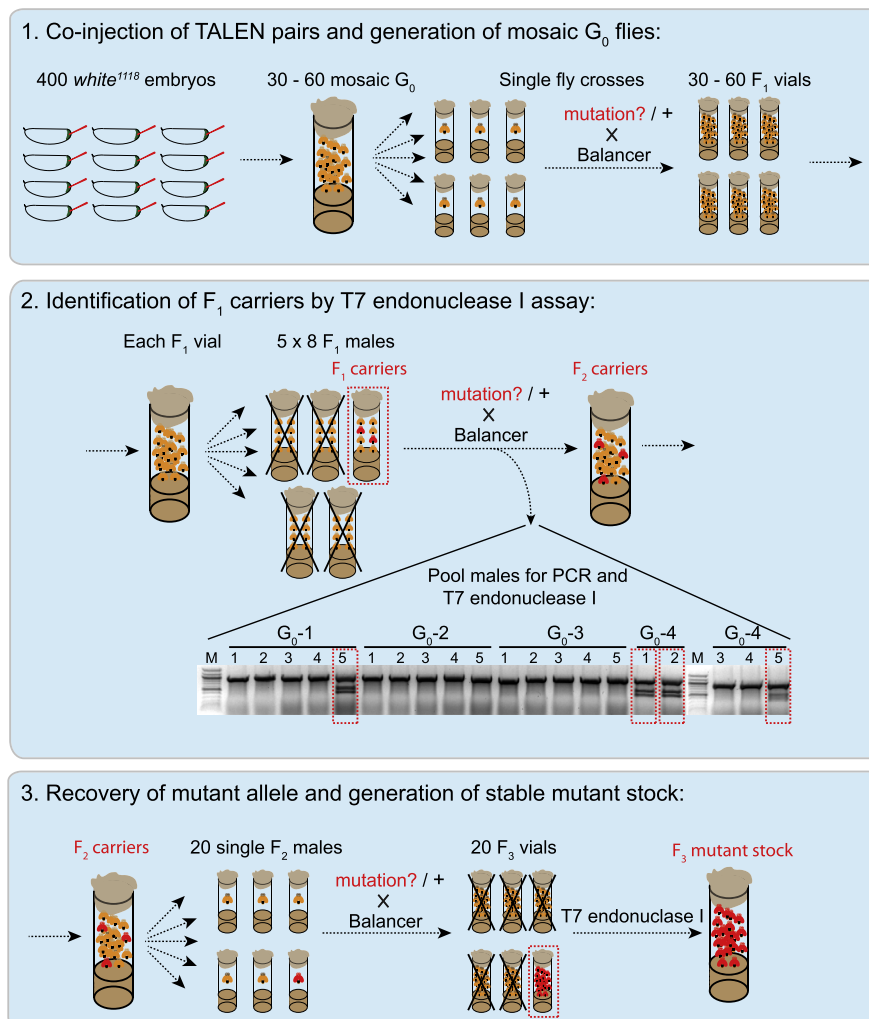


Fig. 2. Detailed overview of fly work and mutant stock establishment. Blue boxes highlight the 3 major steps in the fly work to generate a new TALEN-mediated mutant stock. After assembly, mRNAs for TALEN pairs are injected into embryos and mosaic G_0 flies are singly crossed to a balancer line (1). 40 F_1 males are isolated from each G_0 cross, pooled into groups of 8 and crossed to a balancer line. Males are recovered and screened by the T7 endonuclease I assay to keep only F_2 vials with mutant carriers (2). 20 males are selected from positive F_2 vials, singly crossed to a balancer line and rescreened with T7 endonuclease I to isolate F_3 vials carrying the mutation and generate a stable mutant stock (3). An orange colour denotes wild-type, while red denotes mutants.

males or virgins. Using the T7 endonuclease I assay, we identified five F_1 candidate mutant carriers originating from five different G_0 crosses and were able to recover all five mutant candidates as F_3 stocks (Fig. 3-1). DNA sequencing revealed that *CG11617*¹ contains a 14 bp deletion in the spacer region of the first TALEN pair, located in the first coding exon of *CG11617*. This results in a frame shift and thus is predicted to be a genetic null allele (Fig. 3-3). The other four independent *CG11617* alleles all contain exactly the same 11 bp deletion in the spacer region of the second TALEN pair, which is also located within the first coding exon of *CG11617*. These are also predicted to be genetic null alleles. It may initially appear surprising to generate the identical mutation several times; however, similar observations have been made by the application of TALENs in rat embryos [27] or CRISPR in mouse embryos [28].

These results show the value of injecting two TALEN pairs at the same time: even if one TALEN pair is not effective in generating a mutation (a fact that has been observed in systematic studies using TALENs to mutate genes in human cell lines [29]), mutations caused by the second TALEN pair can additionally be isolated from the same batch of injected embryos. This reduces the fly work and in particular the time consuming PCR and T7 endonuclease screening assays. Using this strategy, we have been able to generate

several null alleles for three novel genes. In some cases, the sequence between both TALEN pairs is entirely deleted (X.Z., F.S. unpublished results). Thus, our protocol is very time effective as injection of only 400 embryos reliably results in mutations in the gene of interest. In the future, this approach should allow to efficiently generate mutations in any *Drosophila* gene of interest and with slight modifications of the injection and screening procedure, would also enable insertion of short tags into a target gene to efficiently tag a protein of choice.

A similar protocol can also be applied to modify *Drosophila* genes using the recently developed CRISPR-Cas9 system. Effective generation of mutations in *Drosophila* using a transgenic source for Cas9 as well as for the gRNAs has been recently reported [30]. While being very effective in mutating the target gene, this method requires the generation of transgenic stocks for each of the gRNAs used to target a gene and thus is potentially more time consuming. In addition, CRISPR-Cas9 carries a much higher off-target rate when compared with TALENs [31], which is potentially even higher, when the fly expresses active Cas9 and gRNA in the germ line for several generations. Furthermore, the use of CRISPR has some sequence restrictions [32] compared to TALENs making TALENs a more flexible reagent.

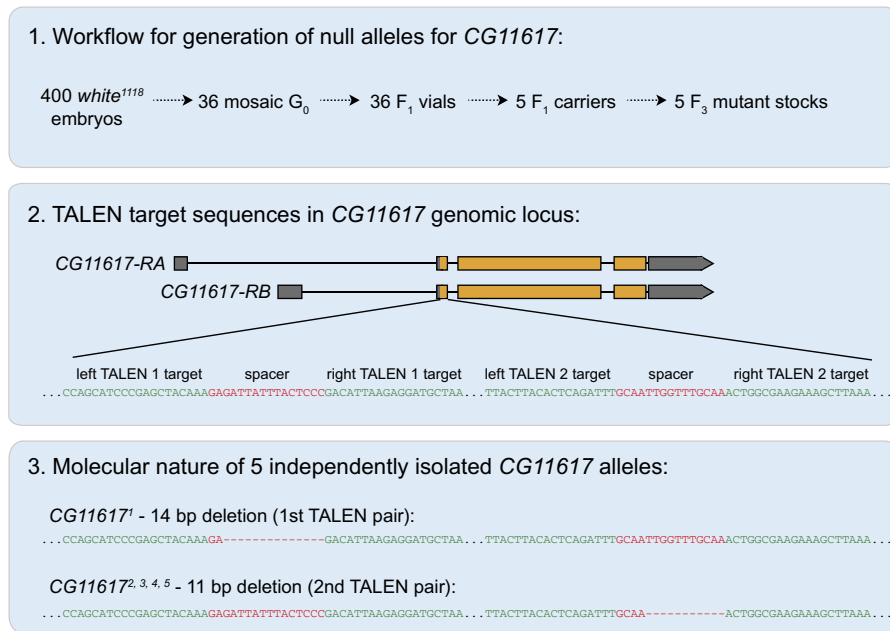


Fig. 3. Generation of null alleles for *CG11617* via TALENs. The workflow to obtain *CG11617* null alleles is illustrated in (1). TALEN target sequences in the *CG11617* locus are illustrated in (2). The molecular nature of the 5 independently isolated *CG11617* mutants is diagrammed in (3). Grey bars represent non-coding sequences and orange bars coding sequences. Green bases denote target sequences. Red bases denote spacer sequence between TALEN recognition sites. Dashes represent missing sequence in mutant alleles.

Acknowledgements

We are very grateful to Bettina Stender and Nicole Plewka for excellent technical assistance. We thank Valeria Soberon for initial help with the Golden Gate system and Maria Spletter for critical comments on the manuscript. This work was supported by the Max Planck Society, the European Research Council and a Career Development Award by the Human Frontier Science Programme to F.S.

References

- [1] C. Nüsslein-Volhard, E. Wieschaus, *Nature* 287 (1980) 795–801.
- [2] C. Nüsslein-Volhard, E. Wieschaus, H. Kluding, *Roux. Arch. Dev. Biol.* 193 (1984) 267–282.
- [3] D. St Johnston, *Nat. Rev. Genet.* 3 (2002) 176–188.
- [4] K.G. Guruharsha, J.-F. Rual, B. Zhai, J. Mintseris, P. Vaidya, N. Vaidya, et al., *Cell* 147 (2011) 690–703.
- [5] G. Dietzl, D. Chen, F. Schnorrer, K.-C. Su, Y. Barinova, M. Fellner, et al., *Nature* 448 (2007) 151–156.
- [6] J.L. Mummery-Widmer, M. Yamazaki, T. Stoege, M. Novatchkova, S. Bhalerao, D. Chen, et al., *Nature* 458 (2009) 987–992.
- [7] F. Schnorrer, C. Schönbauer, C.C.H. Langer, G. Dietzl, M. Novatchkova, K. Schernhuber, et al., *Nature* 464 (2010) 287–291.
- [8] Y.S. Rong, K.G. Golic, *Science* 288 (2000) 2013–2018.
- [9] K.A. Maggert, W.J. Gong, K.G. Golic, *Methods for homologous recombination in Drosophila*, in: *Drosophila*, Humana Press, Totowa, NJ, 2008, pp. 155–174.
- [10] J. Huang, W. Zhou, A.M. Watson, Y.-N. Jan, Y. Hong, *Genetics* 180 (2008) 703–707.
- [11] M. Bibikova, M. Golic, K.G. Golic, D. Carroll, *Genetics* 161 (2002) 1169–1175.
- [12] K.J. Beumer, J.K. Trautman, A. Bozas, J.-L. Liu, J. Rutter, J.G. Gall, et al., *Proc. Nat. Acad. Sci.* 105 (2008) 19821–19826.
- [13] A. Gupta, R.G. Christensen, A.L. Rayla, A. Lakshmanan, G.D. Stormo, S.A. Wolfe, *Nat. Methods* 9 (2012) 588–590.
- [14] A.J. Bogdanove, D.F. Voytas, *Science* 333 (2011) 1843–1846.
- [15] J.K. Joung, J.D. Sander, *Nat. Rev. Mol. Cell Biol.* 14 (2013) 49–55.
- [16] F.D. Urnov, E.J. Rebar, M.C. Holmes, H.S. Zhang, P.D. Gregory, *Nat. Rev. Genet.* 11 (2010) 636–646.
- [17] T. Cermak, E.L. Doyle, M. Christian, L. Wang, Y. Zhang, C. Schmidt, et al., *Nucleic Acids Res.* 39 (2011) e82.
- [18] D. Hockemeyer, H. Wang, S. Kiani, C.S. Lai, Q. Gao, J.P. Cassady, et al., *Nat. Biotechnol.* 29 (2011) 731–734.
- [19] T.J. Dahlem, K. Hoshijima, M.J. Juryne, D. Gunther, C.G. Starker, A.S. Locke, et al., *PLoS Genet.* 8 (2012) e1002861.
- [20] V.M. Bedell, Y. Wang, J.M. Campbell, T.L. Poshusta, C.G. Starker, R.G. Krug, et al., *Nature* 491 (2012) 114–118.
- [21] C. Mussolino, R. Morbitzer, F. Lutge, N. Dannemann, T. Lahaye, T. Cathomen, *Nucleic Acids Res.* 39 (2011) 9283–9293.
- [22] K.J. Beumer, J.K. Trautman, M. Christian, T.J. Dahlem, C.M. Lake, R.S. Hawley, et al., *G3 (Bethesda)* 3 (2013) 1717–1725.
- [23] J. Liu, C. Li, Z. Yu, P. Huang, H. Wu, C. Wei, et al., *J. Genet. Genomics* 39 (2012) 209–215.
- [24] T. Katsuyama, A. Akamammedov, M. Seimiya, S.C. Hess, C. Sievers, R. Paro, *Nucleic Acids Res.* (2013).
- [25] J.C. Miller, S. Tan, G. Qiao, K.A. Barlow, J. Wang, D.F. Xia, et al., *Nat. Biotechnol.* 29 (2010) 143–148.
- [26] F. Schnorrer, A. Ahlford, D. Chen, L. Milani, A.-C. Syvänen, *Nat. Protoc.* 3 (2008) 1751–1765.
- [27] L. Tesson, C. Usal, S. Ménoret, E. Leung, B.J. Niles, S. Remy, et al., *Nat. Biotechnol.* 29 (2011) 695–696.
- [28] H. Wang, H. Yang, C.S. Shivalila, M.M. Dawlaty, A.W. Cheng, F. Zhang, et al., *Cell* 153 (2013) 910–918.
- [29] D. Reyon, S.Q. Tsai, C. Khayter, J.A. Foden, J.D. Sander, J.K. Joung, *Nat. Biotechnol.* 30 (2012) 460–465.
- [30] S. Kondo, R. Ueda, *Genetics* 195 (2013) 715–721.
- [31] Y. Fu, J.A. Foden, C. Khayter, M.L. Maeder, D. Reyon, J.K. Joung, et al., *Nat. Biotechnol.* (2013) 1–6.
- [32] L. Cong, F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, et al., *Science* 339 (2013) 819–823.

Publication II

A Versatile Two-Step CRISPR- and RMCE-Based Strategy for Efficient Genome Engineering in *Drosophila*

Xu Zhang, Wouter H. Koolhaas, and Frank Schnorrer

Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

Keywords: *Drosophila*, CRISPR/Cas9, homologous recombination, RMCE, muscle

A Versatile Two-Step CRISPR- and RMCE-Based Strategy for Efficient Genome Engineering in *Drosophila*

Xu Zhang, Wouter H. Koolhaas, and Frank Schnorrer¹

Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

ORCID ID: 0000-0002-1628-9895 (X.Z.)

ABSTRACT The development of clustered, regularly interspaced, short palindromic repeats (CRISPR)/CRISPR-associated (Cas) technologies promises a quantum leap in genome engineering of model organisms. However, CRISPR-mediated gene targeting reports in *Drosophila melanogaster* are still restricted to a few genes, use variable experimental conditions, and vary in efficiency, questioning the universal applicability of the method. Here, we developed an efficient two-step strategy to flexibly engineer the fly genome by combining CRISPR with recombinase-mediated cassette exchange (RMCE). In the first step, two sgRNAs, whose activity had been tested in cell culture, were co-injected together with a donor plasmid into transgenic Act5C-Cas9, *Ligase4* mutant embryos and the homologous integration events were identified by eye fluorescence. In the second step, the eye marker was replaced with DNA sequences of choice using RMCE enabling flexible gene modification. We applied this strategy to engineer four different locations in the genome, including a gene on the fourth chromosome, at comparably high efficiencies. Our data suggest that any fly laboratory can engineer their favorite gene for a broad range of applications within approximately 3 months.

KEYWORDS

Drosophila
CRISPR/Cas9
homologous
recombination
RMCE
muscle

Reverse genetics is currently booming with the establishment of TALEN- and CRISPR-mediated genome engineering (Hsu *et al.* 2014; Sander and Joung 2014; Joung and Sander 2013). In particular, the CRISPR/Cas9 technology appears to efficiently and specifically introduce double strand DNA breaks in the genome of the organism, which can then be utilized to either introduce point mutations by error-prone nonhomologous end-joining (NHEJ) or integrate heterologous DNA into the chromosome using the homology-directed repair (HDR) pathway (Sander and Joung 2014). In *Drosophila*, CRISPR-induced NHEJ has mainly been utilized to mutate genes that result in a visible, easily scored phenotype, such as white eyes or yellow body color, or to mutate GFP transgenes (Gratz *et al.* 2013a; Sebo *et al.* 2014; Ren *et al.* 2013; Bassett *et al.* 2013). Mutants in genes with no visible phenotype required PCR screening for

their identification; therefore, high mutagenesis rates are important, which might be difficult to achieve at all positions in the fly genome (Kondo and Ueda 2013; Yu *et al.* 2013, 2014; Gokcezade *et al.* 2014; Port *et al.* 2014). Recently, this bottleneck was addressed by applying CRISPR-induced HDR to insert an attP-site together with a visible marker into the gene of interest (Baena-Lopez *et al.* 2013; Gratz *et al.* 2014; Xue *et al.* 2014). In some cases the visible marker was flanked by FRT or loxP sites allowing its excision to only leave one attP site (and one loxP or FRT site) within the gene. This attP site enables the introduction of any given DNA sequence into the gene of interest (Baena-Lopez *et al.* 2013; Gratz *et al.* 2014; Xue *et al.* 2014). However, the efficiency of reporter integration was rather low (Baena-Lopez *et al.* 2013) and only determined at a single genomic locus (Gratz *et al.* 2014; Xue *et al.* 2014), leaving the general applicability to the *Drosophila* genome an open question. Port *et al.* (2014) reported an alternative CRISPR-mediated HDR strategy by using a transgenic single guide RNA (sgRNA) source and crossing it to a transgenic Cas9 source. This method also enabled targeting of somatic cells in a tissue-specific manner, but it required the generation of new transgenic sgRNA lines for every locus (Port *et al.* 2014).

Here, we have developed a highly flexible two-step genome engineering platform that combines CRISPR-mediated HDR with Φ C31 recombinase-mediated cassette exchange (RMCE). In the first step, CRISPR is applied to integrate a splice acceptor and an SV40 terminator

Copyright © 2014 Zhang *et al.*

doi: 10.1534/g3.114.013979

Manuscript received August 19, 2014; accepted for publication October 11, 2014; published Early Online October 15, 2014.

This is an open-access article distributed under the terms of the Creative Commons Attribution Unported License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supporting information is available online at <http://www.g3journal.org/lookup/suppl/doi:10.1534/g3.114.013979/-/DC1>

¹Corresponding author: Max Planck Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany. E-mail: schnorrer@biochem.mpg.de

together with a 3xP3-dsRed eye reporter. This enables both the efficient identification of the targeted event and the creation of a strong loss of function allele. In the second step, two flanking attP sites are utilized to replace the inserted DNA by any DNA of choice using RMCE, an established standard technology in *Drosophila* (Venken *et al.* 2011). Together, this allows flexible cassette exchange to freely manipulate the gene of interest. We successfully applied this method to four different locations in the genome and efficiently generated several allele variants, including a conditional allele, from a single targeting event. Our streamlined CRISPR/Cas9-based and RMCE-based strategies make it practical to flexibly engineer any *Drosophila* gene of choice for a broad range of applications within approximately 3 months.

MATERIALS AND METHODS

Fly strains and genetics

All fly work, unless otherwise stated, was performed at 25° under standard conditions. The *Lig4*[169] null allele (McVey *et al.* 2004) was obtained from the Bloomington *Drosophila* Stock Center, and *y*[1], *M(Act5c-Cas9, [w+])* in *M(3xP3-RFP.attP)ZH-2A, w[1118]* was a gift from Phillip Port and Simon Bullock before publication (Port *et al.* 2014). Both markers (*w+* and 3xP3-RFP) were removed by crossing to heat-shock-Cre. The *y*[1], *M(Act5C-Cas9)ZH-2A, w[1118]* flies were recombined with *Lig4*[169] to obtain *y*[1], *M(Act5C-Cas9)ZH-2A, w[1118]*, *Lig4*[169].

Cell culture

Drosophila Schneider 2 (S2) cells stably expressing myc-Cas9 from a ubiquitin promoter were a gift from Klaus Förstemann before publication (Böttcher *et al.* 2014). S2 cells were cultured in Schneider's *Drosophila* medium supplemented with 10% fetal calf serum (Life Technologies) and penicillin/streptomycin (GE Healthcare). sgRNA activities were tested by transfecting 1 µg sgRNA per 24 wells into the myc-Cas9 cells using Eugene HD (Promega), followed by DNA extraction and a T7-Endonuclease I assay (see supplied protocol for details).

Plasmids

CC6-U6-gRNA_hsp70-Cas9 plasmid was a gift from Peter Duchek before publication (Gokceade *et al.* 2014). pJET1.2-STOP-dsRed: attP1 and splicing acceptor (SA) were amplified with primers XZ82 and XZ83, SV40 terminator with XZ84 and XZ85, and attP2 with XZ88 and XZ89 from DNA extracted from a MiMIC fly line (Venken *et al.* 2011); 3xP3-dsRed was amplified from a fosmid fly line (Langer *et al.* 2010) with primers XZ86 and XZ87. These fragments were cloned into pLR-HD plasmid by Golden Gate cloning (Cermak *et al.* 2011). This assembled attP1-SA-STOP-SV40-3xP3-dsRed-attP2 cassette was amplified with primers XZ195 and XZ196 and blunt cloned into pJET1.2 to generate pJET1.2-STOP-dsRed. Because this STOP-dsRed cassette is flanked by two BsmBI sites, it can be easily assembled with both homology arms (Figure 3A): each homology arm of approximately 1 kb was amplified from genomic DNA of the target genotype with Phusion polymerase (NEB) and blunt-end cloned into pJET1.2 (CloneJET PCR Cloning Kit, Thermo Scientific). Primers used to amplify the homology arms have a 5' BsmBI site enabling Golden Gate assembly with the STOP-dsRed cassette. All primers used are listed in Supporting Information, Table S1. pBS-donor-backbones pBS-GGAC-TTCT, pBS-GGAC-ATGC, and pBS-CGGA-GTGC were constructed by linearizing pBluescript with *Kpn*I and *Sac*II, followed by amplification with primer pairs XZ150 and XZ151, XZ156 and XZ151, and XZ161 and XZ162, respectively, and re-ligation. The generated pBS-donor-backbones harbor two BsmBI sites for donor plasmid assembly.

pJET1.2-STOP-dsRed, pJET1.2-HA-left, pJET1.2-HA-right, and an appropriate pBS-backbone were assembled to the pBS-donor vector by Golden Gate cloning. attB plasmids FRT-2xTY1-FRT-V5 and 2xTY1-V5 fragments were synthesized as gBlocks (IDT) and cloned into the attB plasmid for all three reading frames. For construction of CC6-U6-gRNA_hsp70-Cas9-sgRNA1,3,4,7 and 9 the CC6-U6-gRNA_hsp70-Cas9 vector was cut with *Bbs*I (NEB) and the annealed sgRNA targeting oligos were cloned into it. The vas-ΦC31(3xP3-EGFP.attB) plasmid was obtained from Johannes Bischof (Bischof *et al.* 2007). The attB site was removed by digestion with *Spe*I, followed by re-ligation.

All plasmids for embryo injections were purified with PureLink HiPure Plasmid Midiprep Kit (Life Technologies). Oligos are listed in Table S1.

sgRNA synthesis

The sgRNA dsDNA template was produced using overlap PCR with a small amount of a common sgRNA scaffold primer, a shorter sgRNA amplification primer, and a sgRNA gene-specific primer that includes the T7 promoter (Figure 3C) (Böttcher *et al.* 2014). All sgRNA primer sequences are listed in Table S1. The PCR product was cleaned by Qiagen MinElute kit (Qiagen). The sgRNAs were transcribed with T7-MEGashortscript Kit (Life Technologies) and purified with MEGAclear Transcription Clean-Up Kit (Life Technologies).

Embryo injection

Preblastoderm embryos of the appropriate genotype were de-chorionated and injected with a FemtoJet apparatus (Eppendorf) using self-pulled glass needles (Harvard Apparatus) under standard conditions at room temperature. Injected embryos were kept for 2 d at 18° and the hatched larvae were collected and grown at 25°. For step 1 injections, pBS-donor plasmid, two sgRNAs, and (optionally) the CC6-U6-gRNA_hsp70-Cas9 plasmid were mixed and diluted in water. *Lig4*[169] embryos were injected with CC6-U6-gRNA_hsp70-Cas9 plasmid (100 ng/µl) and pBS-donor plasmid (500 ng/µl). *y*[1], *M(Act5C-Cas9)ZH-2A, w[1118]*, *Lig4*[169] embryos were injected with both sgRNAs (60–70 ng/µl each) and pBS-donor plasmid (500 ng/µl). For step 2 injections, the attB plasmid (150 ng/µl) was mixed with vasa-ΦC31 plasmid (200 ng/µl).

Immunolabeling of IFMs: Hemi-thoraces of adult *Drosophila* were prepared and stained as described (Weitkunat and Schnorrrer 2014). Rabbit anti-Salm was used at 1:50 (Kühnlein *et al.* 1994), mouse anti-Flag (Sigma), mouse anti-V5 (Abcam), and rhodamine phalloidin (Invitrogen) were all used at 1:500. Nuclei were visualized by embedding in Vectashield plus DAPI (Vector Laboratories) and images were acquired on a Zeiss LSM780 confocal and processed with FIJI and Photoshop.

Detailed *Drosophila* genome engineering protocol by CRISPR-RMCE

1. CRISPR-sgRNA design and donor plasmid cloning ~10 d
 - 1.1 Verify sequence of the planned targeting regions for the sgRNAs in the fly strain used and in the S2 cells by sequencing to identify potential polymorphisms compared with the published sequence.
 - 1.2 For designing the sgRNA targeting sites, choose one of the web tools (Beumer and Carroll 2014). We used an interface designed by the Zhang laboratory (<http://crispr.mit.edu>).
 - 1.3 sgRNA production
 - 1.3.1 Generate the dsDNA template for sgRNA *in vitro* transcription as described by Böttcher *et al.* (2014).

- 1.3.2 Transcribe sgRNA by T7-MEGAscript Kit (AM1354; Life Technologies). Use 150- to 250-ng template for a 20- μ l reaction at 37° overnight.
- 1.3.3 Purify sgRNA by MEGAclear Transcription Clean-Up Kit (Life Technologies). Follow the manufacturer's protocol and, in step 3, add an equal volume of 100% ethanol to the sample.
- 1.3.4 Check the sgRNA integrity on a gel and measure the concentration using a photometer (Nanodrop, Thermo Scientific). The expected yield is 50–100 μ g, which is enough for the S2 cell assay and the fly injections.
- 1.4 sgRNA activity assay in S2 cells
 - 1.4.1 Grow S2 cells in Schneider medium with 10% FCS (Life Technologies) to 5–10 $\times 10^6$ cells/ml at 25°.
 - 1.4.2 Dilute cells to 0.7 $\times 10^6$ /ml and plate 1 ml cells in S2 medium with 10% FCS per well in a 24-well plate for each transfection.
 - 1.4.3 Prepare the transfection mix by diluting 1 μ g sgRNA in 50 μ l serum free medium and 4 μ l Fugene HD mix plus 46 μ l serum free medium, mix both, and incubate for 45 min at room temperature.
 - 1.4.4 Add the Fugene/RNA mix to each well and mix gently by pipetting.
 - 1.4.5 After 48–60 hr at 25°, harvest the cells and extract the genomic DNA by QIAamp DNA mini kit (Qiagen).
 - 1.4.6 For the T7 Endonuclease I assay, amplify an approximately 500-bp fragment, which harbors the sgRNA targeting site with Phusion polymerase (NEB), and denature and anneal the PCR product as described by Zhang *et al.* (2014).
 - 1.4.7 Mix on ice 10 μ l annealed PCR product with 10 μ l T7 Endonuclease I master mix [2 μ l T7 endonuclease I buffer, 0.5 μ l T7 endonuclease I (5 units, NEB) and 7.5 μ l water].
 - 1.4.8 Digest at 37° for 15–20 min using a PCR machine and load on 1.5% agarose gel immediately.
 - 1.4.9 Estimate the efficiency of different sgRNAs by comparing the band intensities of the digested and nondigested bands.
- 1.5 Generation of donor plasmid (can be performed in parallel with steps 1.3 and 1.4 to save time)
 - 1.5.1 Amplify left and right homology arms (approximately 1 kb, start as close to the sgRNA cutting site as possible) with Phusion polymerase (NEB) from the fly strain that is used for HDR and clone them into pJET-1.2 according to the CloneJET PCR Cloning Kit (Thermo Scientific).
 - 1.5.2 Assemble the Golden Gate Cloning reaction [50 ng pBS-backbone, 80 ng pJET1.2-HA-left, 80 ng pJET1.2-HA-right, 80 ng pJET1.2-STOP-dsRed, 1.5 μ l 10x T4 ligation buffer, 1 μ l BsmBI (NEB, R0580), 1 μ l T4 ligase (NEB, M0202) add water to 15 μ l].
 - 1.5.3 Ligate in PCR machine using the following cycles: 15 cycles of 37°, 5 min/16°, 10 min/37°, 15 min/50°, 5 min/80°, and 5 min/4°.
 - 1.5.4 Assemble the Plasmid-safe nuclease reaction (15 μ l ligation reaction 3 μ l 10x Plasmid-safe buffer 1.2 μ l 25 mM ATP 1 μ l Plasmid-safe nuclease (Epicentre) 9.8 μ l water).
 - 1.5.5 Incubate at 37° for 60 min in PCR machine and transform 5–10 μ l in bacteria. Most growing colonies will be correct.
2. Fly step 1 - CRISPR-mediated HDR ~6 wk
 - 2.1 Inject 600–800 *Act5C-Cas9, lig4[169]* embryos with pBS-donor (500 ng / μ l) and two sgRNAs (each 60–70 ng/ μ l, targeting close to the chosen homology arms). Collect at least 50 fertile mosaic G₀ flies.

- 2.2 Cross G₀ flies individually (at least 50 vials) either to *yw* flies or appropriate balancer flies and screen all the F₁ progeny for fluorescent red eyes using a fluorescent binocular (Leica MZ16-FA). Keep track of how many independent G₀ founders lead to how many F₁ carrier flies.
- 2.3 Generate stocks from an individual F1 carrier by crossing to balancer flies resulting in an isogenized stock for the engineered chromosome. Verify the targeting event by PCR and sequencing.

3. Fly step 2: Φ C31-mediated RMCE ~6 wk

- 3.1 Inject a “generic” plasmid generated by Venken *et al.* (2011) or this study or your own custom-made gene-specific attB plasmid (150 ng/ μ l) mixed with vasa- Φ C31 plasmid (200 ng/ μ l) into approximately 200 embryos from an amplified stock generated at 2.3.
- 3.2 Cross G₀ flies individually to an appropriate balancer and screen all F₁ progeny for nonfluorescent eyes using a fluorescent binocular (Leica MZ16-FA).
- 3.3 Generate stocks from an individual F1 carrier by crossing to balancer flies, resulting in an isogenized stock for the engineered chromosome. Verify the correct orientation of the RMCE by PCR (will be ~50%).

RESULTS

Strategy overview

We aimed to develop a versatile and efficient strategy to modify the *Drosophila* genome that would allow various genome modifications such as the introduction of single point mutations, protein tags, exon deletions, or other desired changes in the gene of choice. Despite the suggested higher efficiency of CRISPR/Cas9-induced HDR as compared with Zn-finger-induced or TALEN-induced HDR, the identification of successfully targeted carrier flies is still a limiting step in the process. PCR-based screening or melting curve analysis methods require DNA extraction (Beumer *et al.* 2013a), which can be inconvenient for efficient stock generation. Therefore, we decided to develop a two-step strategy as illustrated in Figure 1, which enables efficient identification of the carrier flies and allows entirely flexible genome engineering. In the first step, we insert a 3xP3-dsRed marker enabling easy identification of the HDR event. A strong splice acceptor, followed by STOP codons and an SV40 polyA terminator, precedes the dsRed cassette. The inserted DNA is flanked by two attP sites in opposite orientations, a strategy that we adapted from the popular MiMIC system (Venken *et al.* 2011). If this cassette is inserted into an intron or replaces an endogenous exon, then it results in truncated mRNA of the targeted gene. Thus, step 1 can be used to create a loss-of-function allele (Figure 1).

In the second step, RMCE is applied to replace the DNA between both attP sites by any DNA of choice, leaving a minimal scar of two attR sites, preferably in introns. RMCE has been used very efficiently in the MiMIC system, demonstrating that attR sites in introns generally do not interfere with gene function (Venken *et al.* 2011). Hence, our strategy enables the generation of various alleles, like a defined point mutation, a tagged allele, an exon replaced by a tag, or a conditional allele, from a single HDR carrier (Figure 1). This strategy should allow flexible editing of any *Drosophila* gene within approximately 3 months (Figure 2).

CRISPR design and cloning

In step 1, we aimed to insert a STOP-3xP3-dsRed cassette flanked by two attP sites using a donor plasmid (Figure 1). Because the same strategy should be applicable to any gene, we established a single-step

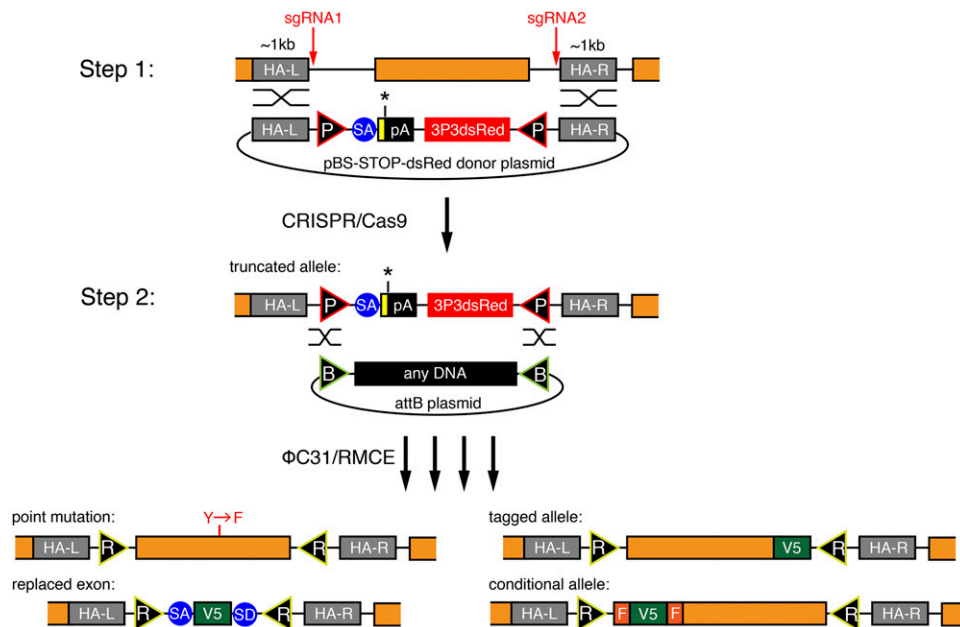


Figure 1 A two-step method to flexibly engineer the fly genome. Overview of the two-step procedure. In step 1, a donor vector consisting of two attP sites (P), a splice acceptor (SA), and STOP codons (yellow box, black asterisk) followed by an SV40 polyadenylation signal (pA) and a 3xP3dsRed marker are inserted on Cas9 cleavage with two sgRNAs. The orange coding exon is excised. In step 2, ΦC31-mediated RMCE inserts any DNA sequence between the two attB sites (B). Examples for various engineered exons are given, resulting in attR sites (R) in introns. F stands for FRT.

Golden Gate protocol to assemble the STOP-dsRed donor plasmid containing approximately 1-kb homology arms on each side, which has been shown to be of sufficient length for efficient HDR (Beumer *et al.* 2013b) and can be easily amplified by PCR. Cloning of the homology arms into the donor vector is thus very straightforward and takes only a few days for the gene of choice (Figure 3A, see *Materials and Methods*). This donor vector is the template for the HDR in step 1.

In step 2, RMCE exchanges the STOP-dsRed by any sequence located between two attB-sites in a provided donor plasmid (Figure 1). RMCE works very reliably and a large collection of plasmids to tag

genes or insert reporters for various applications is available (Venken *et al.* 2011). These plasmids are fully compatible with our step 2 design. We have generated additional “generic” attB plasmids that can be used to tag any gene with a 2xTY1-V5 tag or to engineer a conditional allele using an FRT flanked 2xTY1 cassette followed by a V5 tag (Figure 3B). Flp-mediated deletion of the 2xTY1 cassette will lead to a frame shift and thus can be used to create loss-of-function clones at very high efficiency, as the flip-out will occur *in cis* (Hadjieconomou *et al.* 2011). The TY1 tag is a convenient affinity tag (Sarov *et al.* 2012). We have generated both constructs in all three reading frames.

CRISPR- and RMCE-assisted genome engineering timeline

1. CRISPR design and cloning ~ 10 days

- | | |
|--|----------|
| 1.1 Targeting strategy and sequencing of CRISPR target area in used fly strain | - 2 days |
| 1.2 sgRNA design and cloning of 1 kb donor homology arms into pJET | - 2 days |
| 1.3 sgRNA in vitro transcription and Golden Gate assembly of donor plasmid | - 2 days |
| 1.4 sgRNA activity assay in S2-cells | - 3 days |

2. Fly step 1 - CRISPR-mediated HDR ~ 6 weeks

- | | |
|--|-----------|
| 2.1 sgRNAs and donor vector injection into 600 - 800 <i>Act5C-Cas9</i> , <i>Lig4[169]</i> embryos resulting in more than 50 fertile mosaic G_0 flies | - 12 days |
| 2.2 G_0 fly crosses to balancer and screen F_1 progeny for fluorescent red eyes | - 12 days |
| 2.3 Verification of positive F_1 s by PCR and sequencing, and generation of stocks | - 12 days |

3. Fly step 2 - ΦC31-mediated RMCE ~ 6 weeks

- | | |
|--|-----------|
| 3.1 Injection of “generic” or gene-specific attB plasmid together with vasΦC31 plasmid into ~ 200 embryos from stocks generated at 2.3 | - 12 days |
| 3.2 G_0 fly crosses to balancer and screen F_1 progeny for non-fluorescent eyes | - 12 days |
| 3.3 Verification of positive F_1 s by PCR and generation of final stocks | - 12 days |

Figure 2 Two-step genome engineering timeline. Schematic overview of the major steps of the genome engineering procedure. Details are provided in *Materials and Methods*.

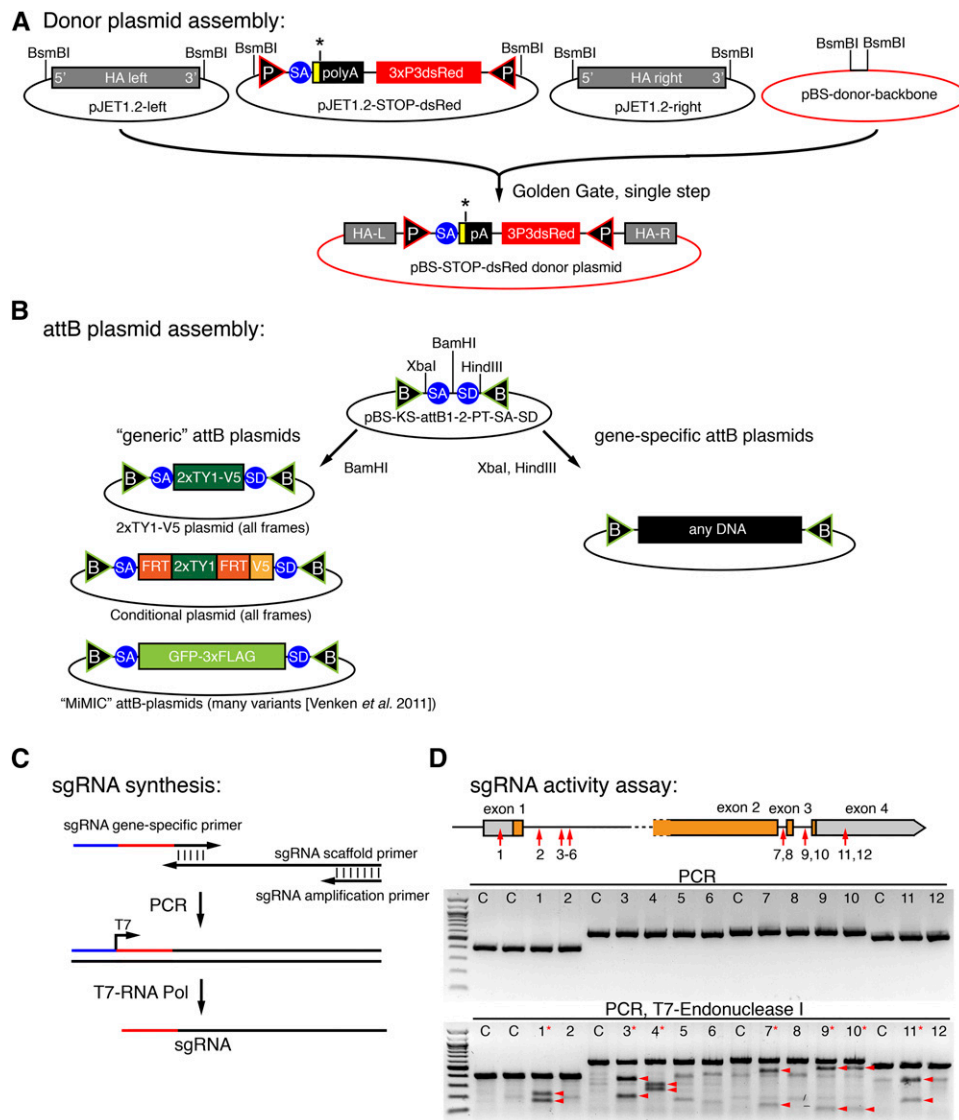


Figure 3 Cloning scheme and sgRNA activity tests. **(A)** Single-step Golden Gate assembly scheme of the STOP-dsRed donor vector cloned into a modified pBluescript backbone. **(B)** Scheme of the “generic” attB plasmids used in our study. A simple cloning step is sufficient to generate any gene-specific attB plasmid that can be used to replace an excised exon. **(C)** sgRNA synthesis scheme. **(D)** sgRNA activity assay of 12 different sgRNAs in S2 cells. PCR result with (bottom) or without (middle) T7-Endonuclease I treatment are shown. Digested products are marked by arrow heads and effective sgRNAs are marked by a red asterisk. C are controls without sgRNA.

CRISPR activity assay in cell culture

Many search algorithms exist to predict sgRNA target sequences for a given gene region (Beumer and Carroll 2014). However, to date there is no simple way of confirming if any of the predicted sgRNAs work efficiently. We developed such a selection assay to be able to only inject effective sgRNAs into fly embryos. We designed 12 different sgRNAs targeting different regions in the *salm* gene and synthesized the sgRNAs by a standard PCR and *in vitro* transcription reaction (Figure 3C). These sgRNAs were then individually transfected into Cas9 expressing S2 cells (Böttcher *et al.* 2014), and their cleavage efficiency was determined with a simple T7-Endonuclease I assay (see *Materials and Methods*) (Zhang *et al.* 2014). On average, approximately half of the tested sgRNAs work efficiently in this assay (Figure 3D), strongly suggesting that such a pre-selection test is useful to improve the *in vivo* success rates.

Step 1: HDR in *Lig4* mutant embryos

To test the efficiency of inserting our STOP-dsRed cassette, we designed three donor constructs targeting different regions in the *salm* gene (Figure 4A): the first, deleting parts of exon 1; the second, inserting the cassette into intron 1; and the third, deleting exon 3. For each construct approximately 1-kb homology arms were cloned in the

STOP-dsRed donor vector. We injected the STOP-dsRed donor as circular plasmid together with two plasmids each containing a U6 promoter-driven sgRNA verified in S2 cells and a hsp70-Cas9 source (see *Materials and Methods*). We injected into *Ligase4* mutant embryos, which were reported to exhibit a higher rate of HDR than wild-type embryos (Beumer *et al.* 2013a,b). We injected between 700 and 1500 embryos for each of the three constructs and were able to recover 11 red-eyed F1 carriers from two independent founders for the first intron construct and 72 red-eyed F1s from four independent founders for the third exon deletion construct (Table 1). This demonstrated that our strategy works in principle, but because we failed to recover the first exon deletion allele, we wanted to further improve the efficiency by using a different Cas9 source.

Step 1: Transgenic Cas9 improves HDR efficiency

A number of transgenic Cas9 flies have been generated recently, and some of which have been used successfully (Ren *et al.* 2013; Xue *et al.* 2014; Gratz *et al.* 2014; Port *et al.* 2014). To test whether a transgenic Cas9 source is more efficient for HDR than a source from an injected plasmid, we targeted the same positions as above, but now using *Act5C-Cas9*, *Lig4* flies. For this, we recombined an *Act5C-Cas9* transgene

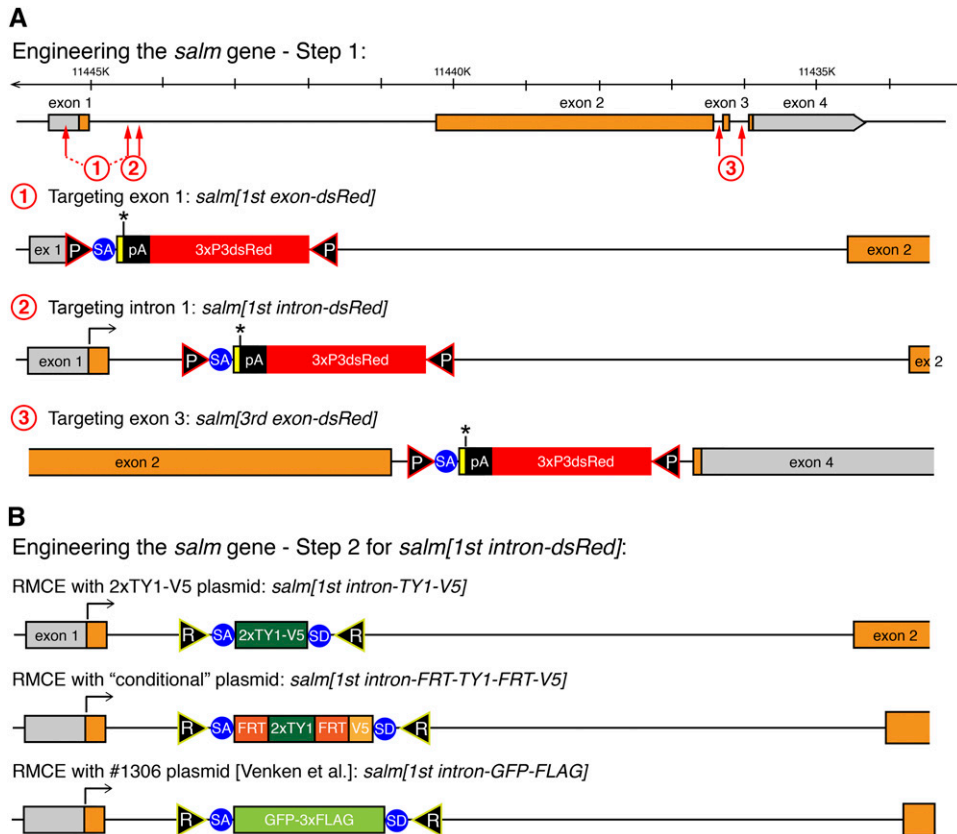


Figure 4 Engineering of the *salm* gene. (A) Step 1 engineering of the *salm* gene. The genomic *salm* organization is depicted with coding exons in orange. The sgRNA targeting sites are indicated by red arrows and the resulting *salm*[1st exon-dsRed], [1st intron-dsRed] and [3rd exon-dsRed] alleles are shown. (B) Step 2 engineering of the *salm* gene. RMCE products of the *salm*[1st intron-dsRed] with three different exon cassettes are shown.

expressing Cas9 ubiquitously, including maternally in the germline (Port *et al.* 2014), with the *Lig4*[169] null allele. Additionally, we removed the *white* and *3xP3dsRed* markers from the *Act5C-Cas9* transgene to obtain a *Act5C-Cas9*, *Lig4*[169] chromosome that is useful for the injection of our donor plasmids (see *Materials and Methods*). We injected 700 *Act5C-Cas9*, *Lig4*[169] embryos with two *in vitro* transcribed sgRNAs targeting either the first exon, the first intron, or the third exon of *salm* (using the same sgRNA target sequences as used above). We obtained one, nine, and four independent founders producing 13, 54, and 59 F1 carriers, respectively, demonstrating that all three locations were targeted successfully with frequencies between 3% and 14% per fertile G₀ (Table 1). To verify that the targeted insertion occurred correctly, we tested a total of nine independent carriers from the three locations by PCR and sequencing. We were able to confirm that all of these targeted correctly by "ends-out" homologous recombination. We did not detect any "ends-in" insertions, which were reported to occur occasionally (Yu *et al.* 2014) (Figure S1 shows *salm*[1st intron-dsRed] as an example).

The first step of our gene-targeting strategy inserts a strong splice acceptor followed by a STOP cassette into the gene and thus should terminate transcription at this position. By design, the *salm*[1st exon-dsRed] allele additionally has a deleted ATG. As expected, the *salm*[1st exon-dsRed] allele is homozygous lethal, as well as lethal *in trans* to *salm*[1], demonstrating that we created a strong *salm* loss-of-function allele (Figure 5A). The *salm*[1st intron-dsRed] allele harbors an insertion in the first intron (Figure 4A). This allele is also homozygous lethal, and lethal *in trans* to *salm*[1], suggesting that the splice acceptor and STOP cassette are used efficiently to create a strong loss-of-function allele (Figure 5A). The *salm*[3rd exon-dsRed] allele only deletes the last 36 amino acids of the long SalmPA isoform, including 10 amino acids of the last zinc finger (Figure 4A). This allele is homozygous viable (Figure 5A). Taken together, these data suggest that our CRISPR-mediated step 1 strategy works efficiently to isolate targeted carrier flies at a practical frequency for routine use. Conveniently, these step 1 alleles are generally loss-of-function alleles if the insertion is located within the gene.

Table 1 Summary of the transformation efficiencies for the four different genomic locations modified in this study

Location	Injected Genotype	sgRNAs	Injected Embryos	G0 Larvae	G0 Adults	Fertile G0 Adults	Independent G0 Founders	Total Red-Eyed F1	Independent Founders Per Fertile G0s
<i>salm</i> [1 st exon-dsRed]	<i>Lig4</i> ^{-/-}	sgRNA1 + sgRNA3	1500	396	166	ND	0	0	ND
<i>salm</i> [1 st intron-dsRed]	<i>Lig4</i> ^{-/-}	sgRNA3	700	171	149	ND	2	11	ND
<i>salm</i> [3 rd exon-dsRed]	<i>Lig4</i> ^{-/-}	sgRNA7 + sgRNA9	1500	312	150	ND	4	72	ND
<i>salm</i> [1 st exon-dsRed]	<i>Act5C-Cas9</i> , <i>Lig4</i> ^{-/-}	sgRNA1 + sgRNA3	700	124	48	30	1	13	3.3%
<i>salm</i> [1 st intron-dsRed]	<i>Act5C-Cas9</i> , <i>Lig4</i> ^{-/-}	sgRNA3 + sgRNA4	700	200	122	64	9	54	14.1%
<i>salm</i> [3 rd exon-dsRed]	<i>Act5C-Cas9</i> , <i>Lig4</i> ^{-/-}	sgRNA7 + sgRNA9	700	291	150	99	4	59	4.0%
<i>bent</i>	<i>Act5C-Cas9</i> , <i>Lig4</i> ^{-/-}	<i>bent</i> -sg1 + <i>bent</i> -sg3	700	204	118	56	1	5	1.8%

Step 2: Flexible gene editing by RMCE

A major benefit of our editing strategy is the flexible step 2 that enables the near-seamless insertion of any DNA sequence with only two remaining attR sites (Figure 1). To test the feasibility of step 2, we chose the *salm*[1st intron-*dsRed*] allele. We exchanged the STOP-*dsRed* cassette with a short 2xTY1-V5 exon, a FRT-2xTY1-FRT-V5 conditional exon, and a large GFP-3xFLAG exon from Venken *et al.* (2011) (Figure 4B). As expected, in all three cases the cassette exchange worked routinely and, typically, injection of approximately 200 embryos is sufficient to obtain two or more RMCE events in the correct orientation (see *Materials and Methods*). Importantly, the *salm*[1st intron-*dsRed*] lethality was reverted by RMCE in all three cases (Figure 5B). This demonstrates that our editing protocol generally does not result in any unwanted lethal mutations on the edited chromosome.

Salm protein is expressed in indirect flight muscles (IFMs) and is essential for fibrillar IFM fate specification (Schönbauer *et al.* 2011). Thus, we should detect the tagged Salm protein versions in the IFM nuclei of adult flies. Tagged protein from all three alleles, *salm*[1st intron-TY1-V5], *salm*[1st intron-FRT-TY1-FRT-V5], and *salm*[1st intron-GFP-

FLAG] is expressed in IFMs. Salm-TY1-V5 and Salm-GFP-FLAG are readily detected in the IFM nuclei, and Salm-FRT-TY1-FRT-V5 shows an additional dotted pattern in the cytosol, which might be caused by the FRT sequence translated into protein (Figure 5C–G). The fibrillar IFM morphology is normal in all three homozygous *salm* alleles, showing that the tagged Salm proteins are indeed functional. Each IFM fiber contains several hundred nuclei. The conditional *salm*[1st intron-FRT-TY1-FRT-V5] should now enable a clonal loss-of-function analysis of *salm* in muscle only, as flip-out *in cis* is highly efficient (Hadjieconomou *et al.* 2011). Thus, this strategy should generally be versatile for the genetic analysis of muscle in the future.

Gene editing on the fourth chromosome

To demonstrate the general applicability of our gene editing strategy, we decided to apply it to an additional locus. We chose the *bent* gene, located on chromosome four, which is highly heterochromatic and thus difficult to manipulate by standard genetic tools. To our knowledge, there is only a single case reported in the literature that targeted a gene located on the fourth chromosome by classical ends-out mediated

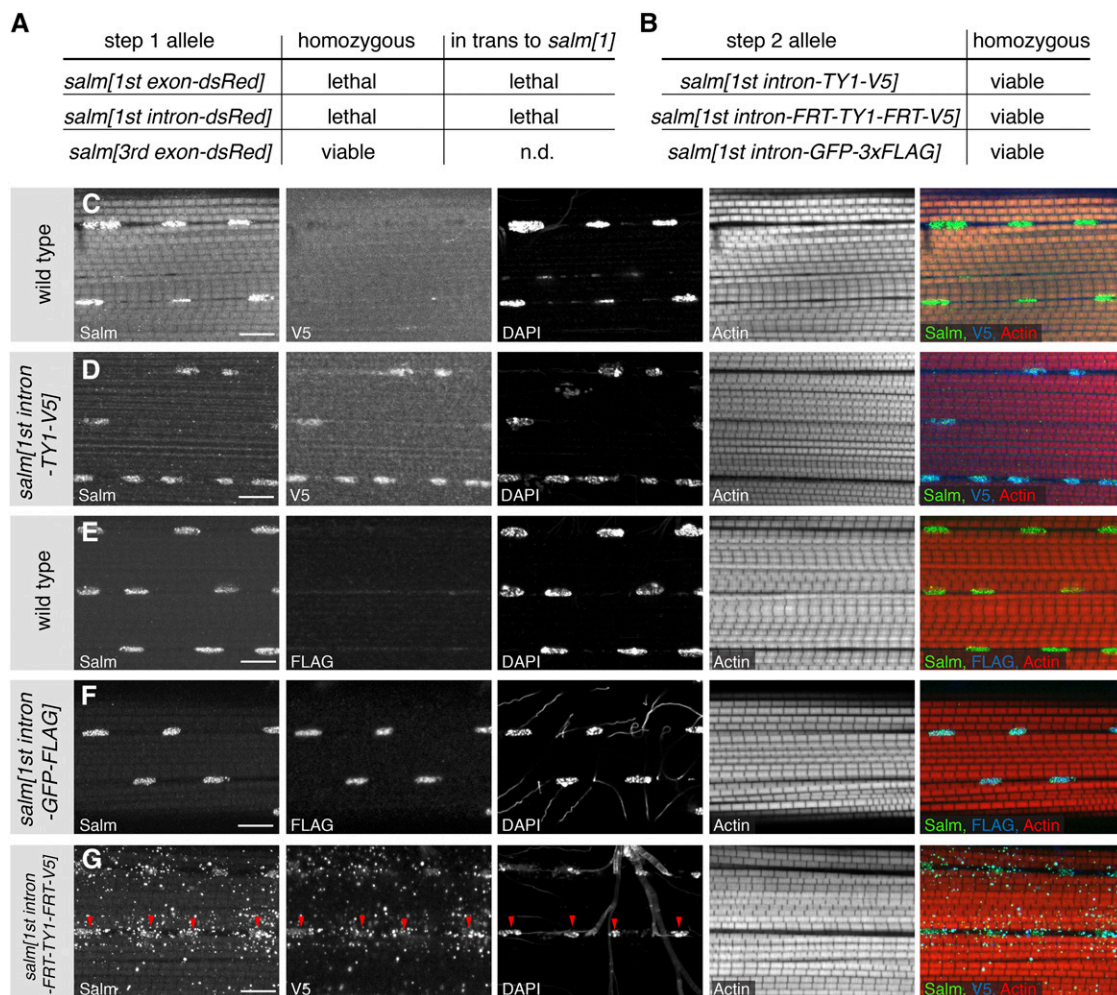


Figure 5 Phenotypic analysis of the engineered *salm* alleles. **(A)** Lethality assay of the *salm*-*dsRed* alleles as homozygous or in trans to *salm*[1]. **(B)** All tagged *salm*[1st intron] alleles regain homozygous viability after step 2. **(C–G)** Localization of the tagged Salm proteins. Untagged Salm is located in the nucleus of wild-type IFMs (C, E), whereas V5 tagged Salm is only detected in the *salm*[1st intron-TY1-V5] and the *salm*[1st intron-FRT-TY1-FRT-V5] alleles (D, G). FLAG is found in the IFM nuclei of *salm*[1st intron-GFP-FLAG] adults (F). The Salm-FRT-TY1-FRT-V5 protein is found in the nuclei (red arrow heads) and also located in dots in the cytoplasm (G). Note the normal fibrillar morphology of the myofibrils in all the homozygous *salm*[1st intron] alleles (D, F, G). Actin was stained with phalloidin and the scale bars are 10 μ m.

homologous recombination using long homology arms (Rodriguez-Jato *et al.* 2011). *bent* is a very large gene composed of at least 46 exons that are spread across more than 51 kb of genomic DNA (Figure 6). *bent* encodes for Projectin, a titin-like protein that is specifically expressed in muscles and essential for correct sarcomeric organization (Fyrberg *et al.* 1992; Ayme-Southgate *et al.* 1995; Schnorrer *et al.* 2010). It is supposedly silent in germ cells, in which the targeting event must happen. We chose to delete exon 11, an exon at the beginning of the PEVK domain of Projectin (Ayme-Southgate and Southgate 2006), using two sgRNAs flanking the exon. Both sgRNAs tested positively in the S2 cell assay (data not shown). We again used approximately 1 kb homology arms and injected the donor vector into 700 *Act5C-Cas9*, *Lig4*[169] embryos. We isolated five carriers from 1 founder out of a total of 56 fertile G₀ flies, resulting in an HDR efficiency of 1.8%. We confirmed the *bt*[11th intron-dsRed] allele by sequencing of the locus. As expected, *bt*[11th intron-dsRed] is homozygous lethal and also lethal *in trans* to *bt*[I-b], a strong bent allele (Ayme-Southgate *et al.* 1995), again suggesting that the inserted splice acceptor is used effectively and transcription is prematurely terminated. Together, these results demonstrate that our CRISPR-mediated targeting strategy also works efficiently on the fourth chromosome, suggesting it can be generally applied to any locus of choice in the fly genome.

DISCUSSION

CRISPR/Cas9 has been used successfully in many model organisms to generate mutants or to introduce targeted changes by HDR (Hsu *et al.* 2014). In *Drosophila*, there has been no general agreement regarding which strategy works most effectively to engineer the genome. To simply mutate a gene by CRISPR/Cas9-induced NHEJ, Cas9 was either injected as mRNA (Bassett *et al.* 2013; Yu *et al.* 2013), provided from an injected plasmid (Gratz *et al.* 2013a; Baena-Lopez *et al.* 2013), or provided from a transgenic source (Kondo and Ueda 2013; Ren *et al.* 2013; Sebo *et al.* 2014). Similarly, the sgRNA was either injected as *in vitro* transcribed sgRNA or provided by an injected plasmid or a transgenic source. A standard protocol has not yet emerged, although several genes have been mutated.

NHEJ can only induce small insertions or deletions. In contrast, HDR allows the defined engineering of a given gene and thus is suitable for a much wider range of applications. CRISPR/Cas9-mediated HDR has been used in *Drosophila* to insert short attP or tag sequences from single-strand oligonucleotides as donors (Gratz *et al.* 2013a) or larger cassettes including a dsRed marker cassette from a plasmid donor (Baena-Lopez *et al.* 2013; Yu *et al.* 2014; Gratz *et al.* 2014; Xue *et al.* 2014), again using various ways of injected or transgenic sources of

sgRNAs or Cas9. The injected genotype was variable; sometimes *Lig4* mutants were used (Baena-Lopez *et al.* 2013; Yu *et al.* 2014; Gratz *et al.* 2014; Xue *et al.* 2014), sometimes they were not used (Baena-Lopez *et al.* 2013; Yu *et al.* 2014; Gratz *et al.* 2014; Xue *et al.* 2014; Gokcezaade *et al.* 2014; Port *et al.* 2014). Often the detection of the targeted event required laborious fly screening by PCR (Gratz *et al.* 2013b; Yu *et al.* 2014; Gokcezaade *et al.* 2014).

Here we aimed to develop a universal and efficient CRISPR-based strategy that enables flexible genome engineering, including the insertion of large tags into the coding region of a gene or the generation of conditional alleles. This strategy should be generally applicable to most *Drosophila* genes. Our results confirmed that approximately 1-kb homology arms are of sufficient length to insert a large marker cassette, as has been suggested before for other loci (Gratz *et al.* 2014; Yu *et al.* 2014; Xue *et al.* 2014; Port *et al.* 2014). Thus, we could develop an efficient donor plasmid assembly protocol that facilitates cloning of the donor vector for any gene within a few days. Additionally, our data support the value of a quick pretesting strategy of predicted sgRNAs in S2 cells to eliminate inefficient sgRNAs, which would likely reduce targeting efficiency *in vivo*. However, we have not tested how well sgRNA efficiencies in S2 cells correlate with efficiencies *in vivo*. Conveniently, the same *in vitro* transcribed RNAs can be used for both S2 cell transfections and embryo injections. Our results suggest that a transgenic Cas9 source mediates HDR effectively in *Ligase4* mutant germline cells. Although *Act5C-Cas9* expression is not restricted to the germline, injections of the donor vector together with two verified sgRNAs led to a targeting efficiency of 2%–14% of fertile G₀ flies for the incorporation of the large STOP-dsRed cassette, even for the *bent* locus on the heterochromatic fourth chromosome. This suggests that approximately 50–100 fertile G₀ flies should be sufficient in most cases to identify positive carriers. We and others (Gokcezaade *et al.* 2014) have observed relatively high developmental lethality of the injected G₀ flies, which might be caused by somatic knock-out of the targeted gene. Thus, survival rate of the injected embryos and larvae might be increased when using a germline-restricted Cas9 source, such as *nos-Cas9*. However, *nos-Cas9* was reported to be less efficient in germline transmission compared with *Act5C-Cas9* (Port *et al.* 2014). We thus far have deleted up to approximately 1 kb of genomic sequence by HDR. Larger deletions would likely occur at reduced efficiencies; however, the dsRed marker should still make it practical to find them. The straightforward identification of carriers together with our simple cloning scheme should easily facilitate the insertion of the STOP-dsRed cassette into the gene of choice.

Recent reports using a transgenic sgRNA source (Port *et al.* 2014) or an injected sgRNA source (Ren *et al.* 2014) report very effective

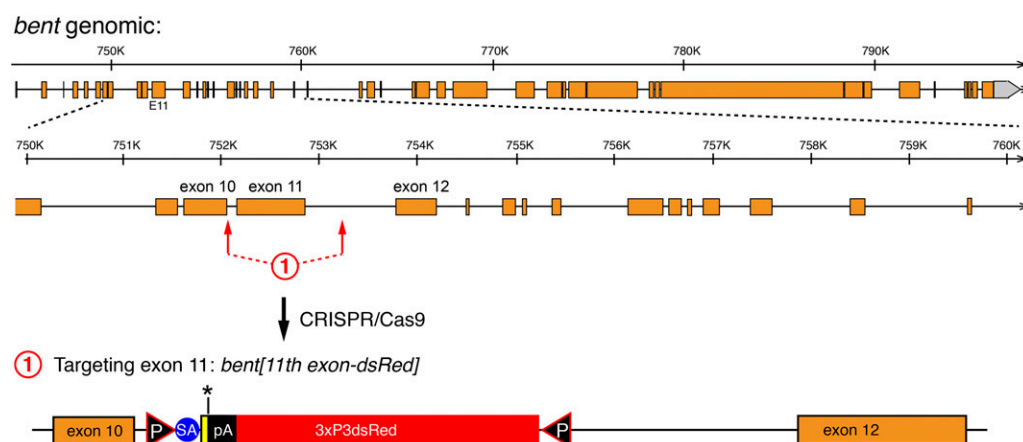


Figure 6 Engineering of the *bent* gene. The entire genomic *bent* organization is shown at the top with a 10-kb zoom-in below. Coding exons are in orange. The sgRNA targeting sites flanking exon 11 are indicated by red arrows and the resulting *bent* [11th exon-dsRed] allele is shown at the bottom.

HDR rates with more than 50% of the fertile *G₀* flies being positive founders and, thus, screening by PCR-based methods were practical in these cases. However, both studies used only a single locus to insert the GFP tagging cassette; hence, a direct comparison with the efficiencies that we report here is difficult.

Our two-step strategy combines the advantages of both CRISPR and RMCE, thus allowing very flexible modifications of a particular gene region with minimal effort. Multiple fluorescent and affinity tags can be easily inserted or a deleted exon can effectively be replaced by various engineered exon versions. In principle, larger gene parts consisting of multiple exons can also be deleted and replaced by modified versions. This method is particularly valuable for genes that harbor complex transcriptional control and function in many tissues such as *salm* or for genes that are exceptionally large and exhibit complex alternative splicing patterns such as *bent*. The two-step strategy allows structure–function analysis at the endogenous locus without interfering with the regulatory regions included in introns, which cannot be achieved by simply inserting a cDNA at the transcriptional start site. The functionality of our method was verified by the reversion of the lethality for the step 2 alleles in the first intron of *salm*. This furthermore suggests that both steps do not generate additional unintended changes on the chromosome. Therefore, we hope that our strategy will promote the wide application of CRISPR-mediated HDR in *Drosophila*, making it a routine tool used in every fly laboratory like EMS mutagenesis or *P*-element-mediated transformation was in the past century.

ACKNOWLEDGMENTS

We thank Klaus Förstemann for the Cas9 expressing S2 cells, Phillip Port and Simon Bullock for the *Act5C-Cas9* flies, Peter Duchek for the CC6-U6-gRNA_hsp70-Cas9 plasmid, Johannes Bischof for the vasa-ΦC31 plasmid, and the Bloomington *Drosophila* Stock Center for fly strains. We are grateful to Reinhard Fässler for generous support and to Bettina Stender, Nicole Plewka, and Christiane Barz for excellent technical assistance. We thank Cornelia Schönbauer and Maria Spletter for critical comments regarding this manuscript. Our work was supported by the Max Planck Society, a Career Development Award from the Human Frontier Science Program (F.S.), the EMBO Young Investigator Program (F.S.), and the European Research Council under the European Union's Seventh Framework Program (FP/2007-2013)/ERC Grant 310939.

LITERATURE CITED

- Ayme-Southgate, A., and R. Southgate, 2006 Projectin, the elastic protein of the C-filaments, pp. 167–176 in *Nature's Versatile Engine: Insect Flight Muscle Inside and Out*, edited by J. Vigoreaux Landes Bioscience, Georgetown.
- Ayme-Southgate, A., R. Southgate, J. Saide, G. M. Benian, and M. L. Pardue, 1995 Both synchronous and asynchronous muscle isoforms of projectin (the *Drosophila* bent locus product) contain functional kinase domains. *J. Cell Biol.* 128: 393–403.
- Baena-Lopez, L. A., C. Alexandre, A. Mitchell, L. Pasakarnis, and J. P. Vincent, 2013 Accelerated homologous recombination and subsequent genome modification in *Drosophila*. *Development* 140: 4818–4825.
- Bassett, A. R., C. Tibbit, C. P. Ponting, and J.-L. Liu, 2013 Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Reports* 4: 220–228.
- Beumer, K. J., and D. Carroll, 2014 Targeted genome engineering techniques in *Drosophila*. *Methods* 68: 29–37.
- Beumer, K. J., J. K. Trautman, M. Christian, T. J. Dahlem, C. M. Lake *et al.*, 2013a Comparing zinc finger nucleases and transcription activator-like effector nucleases for gene targeting in *Drosophila*. *G3 (Bethesda)* 3: 1717–1725.
- Beumer, K. J., J. K. Trautman, K. Mukherjee, and D. Carroll, 2013b Donor DNA utilization during gene targeting with zinc-finger nucleases. *G3 (Bethesda)* 3: 657–664.
- Bischof, J., R. Maeda, M. Hediger, F. Karch, and K. Basler, 2007 An optimized transgenesis system for *Drosophila* using germ-line-specific ΦC31 integrases. *Proc. Natl. Acad. Sci. USA* 104: 3312.
- Böttcher, R., M. Hollmann, K. Merk, V. Nitschko, C. Obermaier *et al.*, 2014 Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured *Drosophila* cells. *Nucleic Acids Res.* 42: e89.
- Cermak, T., E. L. Doyle, M. Christian, L. Wang, Y. Zhang *et al.*, 2011 Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 39: e82.
- Fyrberg, C. C., S. Labeit, B. Bullard, K. Leonard, and E. Fyrberg, 1992 *Drosophila* projectin: relatedness to titin and twitchin and correlation with lethal(4) 102 Cda and bent-dominant mutants. *Proc. Biol. Sci.* 249: 33–40.
- Gokcezade, J., G. Sienski, and P. Duchek, 2014 Efficient CRISPR/Cas9 plasmids for rapid and versatile genome editing in *Drosophila*. *G3 (Bethesda)* 4: 2279–2282.
- Gratz, S. J., A. M. Cummings, J. N. Nguyen, D. C. Hamm, L. K. Donohue *et al.*, 2013a Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* 194: 1029–1035.
- Gratz, S. J., J. Wildonger, M. M. Harrison, and K. M. O'Connor-Giles, 2013b CRISPR/Cas9-mediated genome engineering and the promise of designer flies on demand. *Fly (Austin)* 7: 249–255.
- Gratz, S. J., F. P. Ukken, C. D. Rubinstein, G. Thiede, L. K. Donohue *et al.*, 2014 Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics* 196: 961–971.
- Hadjiconomou, D., S. Rotkopf, C. Alexandre, D. M. Bell, B. J. Dickson *et al.*, 2011 Flybow: genetic multicolor cell labeling for neural circuit analysis in *Drosophila melanogaster*. *Nat. Methods* 8: 260–266.
- Hsu, P. D., E. S. Lander, and F. Zhang, 2014 Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157: 1262–1278.
- Joung, J. K., and J. D. Sander, 2013 TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* 14: 49–55.
- Kondo, S., and R. Ueda, 2013 Highly improved gene targeting by germline-specific Cas9 expression in *Drosophila*. *Genetics* 195: 715–721.
- Kühnlein, R. P., G. Frommer, M. Friedrich, M. Gonzalez-Gaitan, A. Weber *et al.*, 1994 spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the *Drosophila* embryo. *EMBO J.* 13: 168–179.
- Langer, C. C. H., R. K. Ejsmont, C. Schönbauer, F. Schnorrer, and P. Tomancak, 2010 In vivo RNAi rescue in *Drosophila melanogaster* with genomic transgenes from *Drosophila pseudoobscura*. *PLoS ONE* 5: e8928.
- McVey, M., D. Radut, and J. J. Sekelsky, 2004 End-joining repair of double-strand breaks in *Drosophila melanogaster* is largely DNA ligase IV independent. *Genetics* 168: 2067–2076.
- Port, F., H.-M. Chen, T. Lee, and S. L. Bullock, 2014 Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 111: E2967–E2976.
- Ren, X., J. Sun, B. E. Housden, Y. Hu, C. Roesel *et al.*, 2013 Optimized gene editing technology for *Drosophila melanogaster* using germ line-specific Cas9. *Proc. Natl. Acad. Sci. USA* 110: 19012–19017.
- Ren, X., Z. Yang, D. Mao, Z. Chang, H.-H. Qiao *et al.*, 2014 Performance of the Cas9 nickase system in *Drosophila melanogaster*. *G3 (Bethesda)* 4: 1955–1962.
- Rodriguez-Jato, S., A. Busturia, and W. Herr, 2011 *Drosophila melanogaster* dHCF interacts with both PcG and TrxG epigenetic regulators. *PLoS ONE* 6: e27479.
- Sander, J. D., and J. K. Joung, 2014 CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32: 347–355.
- Sarov, M., J. I. Murray, K. Schanze, A. Pozniakovski, W. Niu *et al.*, 2012 A genome-scale resource for in vivo tag-based protein function exploration in *C. elegans*. *Cell* 150: 855–866.
- Schnorrer, F., C. Schönbauer, C. C. H. Langer, G. Dietzl, M. Novatchkova *et al.*, 2010 Systematic genetic analysis of muscle morphogenesis and function in *Drosophila*. *Nature* 464: 287–291.

- Schönbauer, C., J. Distler, N. Jährling, M. Radolf, H.-U. Dodt *et al.*, 2011 Spalt mediates an evolutionarily conserved switch to fibrillar muscle fate in insects. *Nature* 479: 406–409.
- Sebo, Z. L., H. B. Lee, Y. Peng, and Y. Guo, 2014 A simplified and efficient germline-specific CRISPR/Cas9 system for *Drosophila* genomic engineering. *Fly (Austin)* 8: 52–57.
- Venken, K. J. T., K. L. Schulze, N. A. Haelterman, H. Pan, Y. He *et al.*, 2011 MiMIC: a highly versatile transposon insertion resource for engineering *Drosophila melanogaster* genes. *Nat. Methods* 8: 737–743.
- Weitkunat, M., and F. Schnorrer, 2014 A guide to study *Drosophila* muscle biology. *Methods* 68: 2–14.
- Xue, Z., M. Ren, M. Wu, J. Dai, Y. S. Rong *et al.*, 2014 Efficient gene knock-out and knock-in with transgenic Cas9 in *Drosophila*. *G3 (Bethesda)* 4: 925–929.
- Yu, Z., M. Ren, Z. Wang, B. Zhang, Y. S. Rong *et al.*, 2013 Highly efficient genome modifications mediated by CRISPR/Cas9 in *Drosophila*. *Genetics* 195: 289–291.
- Yu, Z., H. Chen, J. Liu, H. Zhang, Y. Yan *et al.*, 2014 Various applications of TALEN- and CRISPR/Cas9-mediated homologous recombination to modify the *Drosophila* genome. *Biol. Open* 3: 271–280.
- Zhang, X., I. R. S. Ferreira, and F. Schnorrer, 2014 A simple TALEN-based protocol for efficient genome-editing in *Drosophila*. *Methods* 69: 32–37.

Communicating editor: B. J. Andrews

Manuscript I

The BTB-Zinc finger *fallen angel* is an essential transcriptional regulator during sarcomere maturation to adjust muscle force production

Xu Zhang¹, Maria M. Spletter^{1,2}, Bianca H. Habermann^{3,4} and Frank Schnorrer^{1,4}

¹ Muscle Dynamics Group, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

² Department of Physiological Chemistry, Biomedical Center, Ludwig-Maximilians-University of Munich, Großhaderner Str. 9, 81377 Munich, Germany

³ Computational Biology Group, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

⁴ Developmental Biology Institute of Marseille (IBDM), CNRS, UMR 7288, Aix-Marseille Université, Case 907, Parc Scientifique de Luminy, 13288 Marseille, France

Correspondence should be addressed to:

frank.schnorrer@univ-amu.fr

keywords: muscle; sarcomere; hypercontraction; BTB-zinc finger; force

Abstract

Skeletal muscles utilize sarcomeres to produce active contractile forces that enables animal movements. Sarcomeres are complex stereotype units with a pseudo-crystalline regularity of actin, myosin and titin filaments. How these regular units are built during development is poorly understood. Here, we study the *Drosophila* flight muscles in order to better understand how the initially 1.7 μm long immature sarcomeres mature to the 3.2 μm long sarcomeres of the adult. We identified the novel BTB-zinc finger *fallen angel* (*fall*, CG32121) as a muscle-specific transcriptional regulator in a genome-wide RNAi screen in *Drosophila*. Fall is specifically expressed during sarcomere maturation in flight muscles and forms nuclear bodies, which localize close to but are distinct from insulator bodies. Interestingly, *fall* mutants show a defect in sarcomere maturation resulting in too short sarcomeres, which produce too high forces. This results in rupturing of the muscle-tendon apparatus during development and a severe muscle atrophy in adult *fall* mutant animals. Mechanistically, we show that Fall regulates the induction of a subset of sarcomeric genes, which are specifically induced during sarcomere maturation. Thus, Fall ensures the correct stoichiometry of sarcomeric components, allowing their pseudo-crystalline assembly in mature contractile sarcomeres. BTB-zinc finger proteins constitute a large conserved protein family and we speculate that a similar defect might be the cause of the fatal congenital contractures found in humans with mutation in *ZBTB42*, underscoring the general importance of correct sarcomere maturation in higher animals.

Introduction

Muscles produce forces to power the controlled movements of all higher animals. The force producing mini-machines of muscle fibers are the sarcomeres. Sarcomeres are stereotyped units of pseudo-crystalline regularity: both ends of bipolar myosin (thick) filaments located in the sarcomere center are facing towards polar actin (thin) filaments, which are cross-linked at defined Z-discs bordering each sarcomere. The two filaments systems are stably connected by flexible titin filaments (Ehler and Gautel, 2008). Both, the molecular components as well as the architectural organization of sarcomeres are conserved from invertebrates to humans (Vigoreaux, 2006), enabling a myosin sliding mechanism to produce force during sarcomere contraction in all species (Huxley and NIEDERGERKE, 1954; Huxley and Hanson, 1954).

Different muscle fiber types can differ in their exact sarcomeric composition (Schiaffino and Reggiani, 2011; Spletter and Schnorrer, 2014), however sarcomere length is precisely controlled for each fiber type, ranging from 3.0 to 3.4 μm in relaxed human muscle *in vivo* (Llewellyn et al., 2008). Hence, both sarcomeric architecture as well as sarcomeric force production must be regulated accurately to build a functional sarcomere during muscle development.

Muscle development is a multi-step process. Myoblasts fuse to form myotubes (Kim et al., 2015), myotubes elongate to attach both ends to tendon cells (Schnorrer and Dickson, 2004; Schweitzer et al., 2010), and finally myotubes convert to myofibers by assembling their myofibrils. Myofibril formation is also a complex and multi-step process that is not well understood in molecular detail (Sanger et al., 2010; Sparrow and Schöck, 2009). It was recently shown that mechanical tension is important to coordinate the simultaneous assembly of early immature myofibrils that span across the muscle fiber (Lemke and Schnorrer, 2016; Weitkunat et al., 2014). These immature myofibrils undergo a complex maturation process resulting in the pseudo-crystalline structure of mature myofibrils and sarcomeres (Sanger et al., 2010; Spletter et al., 2015). Defects in sarcomere maturation can cause muscle hyper-contraction resulting in severe muscle atrophy already during development or within the first days of life in flies (Montana and Littleton, 2004; Spletter et al., 2015) as well as in vertebrates (Johnston et al., 2000; Patel et al., 2014; Robinson et al., 2007). Thus, correct sarcomere and myofibril maturation are critical for correct force production and thus for muscle function.

Appendices

We use the *Drosophila* indirect flight muscles (IFMs) to study the mechanism of myofibril and sarcomere maturation. Early immature myofibrils display a muscle myosin II pattern with a periodicity of about 1.8 μm at 30 h APF and about 1.9 μm at 48 h APF (pupae raised at 27 °C) (Weitkunat et al., 2014). Until eclosion (> 90 h APF at 27 °C) the myofibrils mature to a thickness of about 1.4 μm , and each sarcomere elongates to its final length of 3.2 μm displaying its pseudo-crystalline regularity (Reedy and Beall, 1993; Spletter et al., 2015; Weitkunat et al., 2014). Correct sarcomere morphology and function require the correct stoichiometry of all sarcomeric components. This is most impressively shown by the dominant flightless phenotypes of heterozygous mutations in muscle myosin heavy chain (*Mhc*) {Bernstein:1983vq} or flight muscle specific actin (*Act88F*) (Cripps et al., 1994). Similarly, a mutation in the regulatory region of troponin I (*wupA*), resulting in reduced levels of the troponin complex, cause flightlessness and flight muscle atrophy in adults (Firdaus et al., 2015). Analogously, mutations in a large number of sarcomeric proteins cause dominant cardiomyopathies in humans {Seidman:2001dn}. Hence, correct stoichiometry of the numerous sarcomeric components must be tightly controlled during myofibril formation and maturation.

We have found that the expression of a significant set of sarcomeric components is strongly induced in flight muscles from 48 h APF onwards, thus ensuring correct sarcomere maturation (Spletter, Zhang & Schnorrer, manuscript in preparation). However, apart from the general muscle specific transcriptional regulator Mef2, which continuously required during IFM development (Soler et al., 2012), and the IFM specific transcriptional regulator Spalt major (Salm), which is important for expression of most IFM specific genes and gene isoforms (Schönbauer et al., 2011; Spletter et al., 2015), we know little about the transcriptional regulation of muscle during sarcomere maturation.

Here, we identify an important role for *fallen angel* (*fall*, *CG32121*), a member of the conserved BTB-zinc finger family of transcriptional regulators, to boost transcription of a subset of sarcomeric proteins whereas down-regulating others. Upon mutation of *fall*, an incorrect stoichiometry of sarcomeric proteins in mutant flight muscles leads to uncontrolled force production, muscle hypercontraction and muscle atrophy during stages of normal sarcomere maturation. Thus, *fall* adjusts the muscle-specific transcriptional program to enable correct sarcomere maturation.

Methods and materials

Fly strains and genetics

All fly work was performed at 27 °C under standard conditions unless specified. *CG32121* deficiency line (*w¹¹¹⁸*; *Df(3L)ED4502, P{3'.RS5+3.3}ED4502/TM6C, cu¹ Sb¹*) was obtained from Bloomington stock center. Fall-GFP Fosmid (FF10059) was generated as described (Sarov et al., 2016). Fall-GFP Fosmid (FF10059) was recombined with *fall[1]* or *fall[2]* respectively and the recombinant was identified by PCR for the rescue experiment. The tissue specific rescue was performed at 18 degree by driven *UAS-fall-HA* (obtained from Johannes Bischof (Bischof et al., 2013)) with a IFM specific driver *Actin88F-GAL4* (Bryantsev et al., 2012) in the *fall[1]* or *fall[2]* background. Hypercontraction rescue were performed with *Mhc[10]; fall[1]* or *Mhc[10]; fall[2]* flies by using the IFM specific myosin mutant *Mhc[10]*(Collier et al., 1990). In order to reinforce the resistance of muscle tendon attachment, a constitutively-active form of Rap1 was overexpressed in IFM by driven *UAS-Rap1-GFP-CA* or *UAS-Rap1-HA-CA* (Ellis et al., 2013)(from Guy Tanentzapf) with *Actin88f-GAL4* (Bryantsev et al., 2012) in *fall[1]* background. This cross was done at 18 degree to eliminate the adverse effect caused by the *Actin88f-GAL4* itself. Flies bearing an extra copy of *Pinch-FL* (Pronovost et al., 2013)(from Julie L. Kadrmas) were balanced with *fall[1]* to generate *Pinch-FL; fall[1]*. Tendon cell labeling was performed by crossing *UAS-CD8GFP* with *stripe-GAL4* as control and crossing *UAS-CD8GFP; fall[1]* with *fall[1]*, *stripe-GAL4*, recombinant of *fall[1]* and a tendon cell specific driver *stripe-GAL4* (Fernandes et al., 1996). The indirect flight muscles were labeled with *Mef2-GAL4; UAS-gmaGFP* to facilitate the dissection of samples for RNA-seq and proteomics. Both *CTCF[GE24185]* and *CTCF[P30.6]* (from Rainer Renkawitz) (Mohan et al., 2007) are strong alleles for *CTCF*. *BEAF-32[AB-KO]* {Roy:2007gq} (from Robert J Johnston Jr) is a null allele for both *BEAF-32A* and *BEAF-32B*. *CP190[1]* and *CP190[2]* (Butcher, 2004) were obtained from Jordan Raff.

Generation of *fall[1]* and *fall[2]* alleles

fall[1] was generated as described in published TALEN methods (Zhang et al., 2014a). Briefly, two pairs of TALENs (pair 1 targeting AACCGCAGCATCATC in the first coding exon of *fall* and pair 2 targeting ACTCCGGAGAGGTGA in the second coding exon of *fall*) have been designed and constructed respectively. In *vitro* transcribed TALEN mRNAs were injected into *w¹¹¹⁸* fly embryos at a concentration of 250 ng/μl

each. The genomic region covering both targeted sites have been amplified by PCR with primer XZ13 and XZ14 and then followed by T7 assay to identified successful targeting. Fly stocks with expected mutation have been established and characterised by sequencing afterwards.

fall[2] was constructed by CRPSPR-based genome editing as published with small optimisation (Zhang et al., 2014b). In brief, the intron sequence between coding exon 1 and 2 of *fall* and sequence from the end of 3'UTR of *fall* was chosen for sgRNA designing with CRISPRscan (Moreno-Mateos et al., 2015). Four sgRNAs were picked for targeting each region according to the score from CRISPRscan and their activities were estimated in S2 cells as described (Zhang et al., 2014b). sgRNA 1 (targeting sequences GTGTAATGCGGTGAAAGCG in the intron region) and sgRNA 2 (targeting sequences GGGTCTAAGACGTTGGTTT in the end of 3'UTR) were picked and went for fly embryo injection. The left homology arm around ~2kb and right arm around ~2kb were assembled together with dsRed cassette as described (Zhang et al., 2014b). The *y[1]*, *M(Act5C-Cas9)ZH-2A*, *w[1118]*, *Lig4[169]* embryos were injected with donor plasmid at 800 ng/μl and sgRNAs at 272 ng/μl each. F1 flies with dsRed eyes were chosen for characterization by PCR and sequencing to confirm the insertion position of the selection cassette. Fly stocks with correct insertion were establish as *fall[2]*. For the RMCE exchange, a full length of *fall*-CDS, *fall*-ΔBTB, and *fall*-ΔZFs were tagged with 2xHA and cloned as described (Zhang et al., 2014b). The plasmids were injected to *fall[2]* fly embryos and the progeny were screened for non fluorescent red eyes and characterised by PCR and sequencing as described (Zhang et al., 2014b).

Sarcomere length quantification

Flies samples were staged to desired stage and dissected. The IFM were stained with rhodamine phalloidin and the 40X4 images were acquired with Leica 780. Then the sarcomere length was quantified by a macro that has been integrated into Fiji (written by Giovanni Cardone, unpublished) automatically. Sarcomere length from at least three images (independent flies) was measured for each group of sample. The two-tailed unpaired student t-test were performed between wt and *fall[1]* or *fall[2]*. Error bar represents standard deviation.

mRNA-seq and proteomics analysis

For mRNA-Seq, IFMs were dissected at either 30h APF or 48h APF and isolated based on *Mef2-GAL4*, *UAS-GFP-Gma* expression (Spletter et al., 2015). For the RNA-seq, three replicates were used and each replicate contains indirect flight muscles dissected

from 150 pupae. For the proteomics, the indirect flight muscles were labeled similarly with RNA-seq experiment. Three replicates were used and each replicate consists of indirect flight muscles dissected from 50 pupae.

mRNA-Seq libraries were prepared as described previously (Spletter et al., 2015). Libraries were sequenced on an Illumina HiSeq 2500 and multiplexed 3-4 samples per lane, obtaining ~100 million reads per library. Sequencing was performed at the Vienna Biocenter Core Facility (VBCF) (<http://www.vbcf.ac.at/facilities/next-generation-sequencing/>). Reads were filtered and trimmed using the FASTX Toolkit and cutadapt and mapped to the Flybase 2015_04 genome assembly using STAR {Dobin:2015by}. Reads were visualized on the UCSC server by normalizing to the largest library size. Libraries were evaluated with featureCounts v1.4.2 {Liao:2014cj}, and differential expression analysis was performed with DESeq2 {Love:2014ka} or DEXSeq (Anders et al., 2012). Additional data processing was handled in R, using packages including ggplot2, GOpplot, plyr, reshape2, VennDiagram and Mfuzz. GO analysis was performed with Gorilla (<http://cbl-gorilla.cs.technion.ac.il/>), REVIGO and GOElite (Supek et al., 2011; Zamboni et al., 2012). Gene set enrichment analysis was performed with GSEA {Subramanian:2005jt, Mootha:2003wx} using user-specified gene sets. Genes induced during IFM myofibril maturation are from a manuscript in preparation (Spletter, Zhang & Schnorrer).

The samples for the proteomics were homogenized in 50 µl 6M of guanidinium chloride. Heat the samples at 95 degree for 3-5 min and then centrifuge at 14000g for 10 min. The supernatant was load on LC-MS according to a published protocol (Cox et al., 2014; Eberl et al., 2013). The data were preprocessed by MaxQuant (Cox and Mann, 2008) and then analyzed with Perseus (Tyanova et al., 2016).

Dissection and Immunostaining

All IFM dissections, if not specified, were performed with a modified procedure based on published protocol (Weitkunat and Schnorrer, 2014). Briefly, after removing of the pupal case from the staged samples, two to three holes were pinned in the abdomen with insect pins. The samples were then transferred to 4% PFA in PBST (0.5% Triton-X in PBS) and fixed for ~ 20 min at room temperature (RT) on a belly dancer. Samples were transferred to a silicon coated petri dish and their position was fixed as described (Weitkunat and Schnorrer, 2014). Cut a hole on the head then continue the cutting to the abdomen along the ventral middle line. Other tissues in between of the two thorax halves were removed by forceps or blew away by pipetting. The two thorax halves were then cut

Appendices

off and transferred to 24 well plate. The samples were washed with 3x PBST for 10 min each at RT before blocking with 5% NGS in PBST for 1 hour at RT. Afterwards, they were incubated with primary antibodies at 4 degree overnight and washed 3x in PBST (10 min each) at RT. Incubation with secondary antibodies were performed at RT for 2 hours. The samples were washed again 3x in PBST and embedded in VECTASHIELD plus DAPI (H-1200, VECTEUR LABORATORIES) to visualise nuclei. The samples were stored at 4 degree before imaging. Images were acquired on a Zeiss LSM780 confocal and processed with FIJI (ImageJ).

To stain Su(Hw) and Mod(mdg4)67.2, the samples were dissected first in PBS or PBS + 250mM NaCl (takes about 25-30 min) before fix with methanol. The fixation were done in precooled 100% methanol at -20 °C for 15 min. Gently shake the 24 well plate every 5 min during fixation. Afterwards, equal volume of PBST were added into the fixation solution slowly and mixed well. Then discard and wash the samples as described above before proceed to immunostaining.

To induce osmotic stress, the samples were incubated in PBS + 250 mM NaCl or PBS + 500 mM sucrose at RT for 30 min on a belly dancer before fixation.

Antibodies and dilutions:

For immunostaining, Rabbit anti-GFP (dilution 1:2000) was obtained from Amsbio (catalog number: TP401). Rat monoclonal anti-HA (3F10) and mouse monoclonal anti-HA (16B12) were used at 1:500. Rabbit anti-CP190 (dilution 1:2000), mouse anti-CP190 (dilution 1:1000), rabbit anti-Mod(mdg4)67.2 (dilution 1:1000) and rabbit anti-Su(Hw) (1:1000) were gifts from Professor Mariano Labrador (Schoborg et al., 2013). Mouse anti-BEAF-32 (dilution 1:20) {Blanton:2003gy}, mouse anti- β -PS (CF.6G11, dilution 1:100) (Brower et al., 1984), mouse anti-futsch (22C10, dilution 1:100) {Zipursky:1984un}, mouse anti-laminDm0 (ADL67.10, dilution 1:50) (Riemer et al., 1995), and mouse anti-HP1a (C19A, 1:100) {James:1986vk} were obtained from DSHB. Guinea pig anti-Shot (1:600) (Strumpf and Volk, 1998) was a gift from Talila Volk. All the primary antibodies were diluted in PBST with 5% NGS. All the secondary antibodies, including goat anti-mouse, goat anti-rat, goat anti-pig, and goat anti-rabbit, and phalloidin rhodamine were purchased from Molecular Probes and applied at 1:500 dilution.

For western blot, muscle fibers dissected from 10 pupae with desired stage were homogenized in 50 μ l SDS buffer (250 mM Tris pH 6.8, 30% glycerol, 1% SDS, 500

mM DTT) then boiled at 95 degree for 5 min. The samples were centrifuged at 12000 rpm for 5 min and 15 µl were loaded to each lane of a 4-12% gradient SDS-PAGE gel. Wet-transfer was performed at 300 mA for 30 min. PVDF membrane (Immobilon, Millipore) were blocked with 10% milk powder and then probed with primary antibodies respective antibodies. Rabbit anti-Gelsolin (dilution, 1:2000) {Stella:1994vz} was a gift from Maria Leptin. Rabbit anti-Flightin (dilution, 1:10000) (Reedy et al., 2000) was a gift from Jim Vigoreaux. Same laminD0 (ADL67.10, 1:2000) and HA (16B12, 1:2000) that mentioned above were used for WB. Goat anti-rabbit IgG (111-035-144, 1:20000) and goat anti-mouse (115-034-146, 1:10000) were obtained from Jackson ImmunoResearch. Visualization was performed with chemiluminescence (Millipore) and using a LAS4000 detector system (FujiFilm).

Flight test

Flight tests were done as previously described (Schnorrer et al., 2010). Around 20-30 fresh adult males (1-3 days after eclosion) were collected and recovered for 48 hours in the incubator at 27 degree. Flip the samples into the flight test cylinder that is divided into 5 zones. Record the landing position of the flies immediately after flipping and count the numbers of flies in each zone of the cylinder.

Results:

Mutation of *fallen angel* (*fall*) results in muscle atrophy

We had identified *CG32121* in a genome-wide muscle-specific RNAi screen resulting in flightless animals with missing flight muscles in adults (Schnorrer et al., 2010). *CG32121* is located on the third chromosome and encodes a BTB-zinc finger, for which no classical mutants were available (Figure 1A). We used our established TALEN and CRISPR protocols (Zhang et al., 2014a; 2014b) to generate two new *CG32131* deletion alleles. Both alleles result in viable but entirely flightless animals, hence gave *CG32121* the name *fallen angel* or *fall* for short (Figure 1B). The TALEN induced *fall*[1] allele contains a 363 bp deletion spanning the splice acceptor of the second exon and a large part of the BTB domain. In the CRISPR-induced *fall*[2] allele we exchanged most of the *fall* locus with a STOP cassette and a 3xP3-dsRed marker leaving only 38 amino acids of the original Fall protein (Figure 1A). Hence, we generated two *fall* alleles, with at least *fall*[2] likely being a null allele, confirming that *fall* is essential for flight.

Histological analysis of the flight muscle morphology in *fall* homozygous or trans-heterozygous mutants over a large *fall* deficiency showed that *fall* is required for flight muscle formation or maintenance in young adults (Figure 1E-J). Importantly, this muscle atrophy phenotype could be rescued by re-expressing Fall-GFP under its endogenous control using a GFP-tagged genomic fosmid construct (Sarov et al., 2016) (Figure 1K-M), demonstrating that the *fall* loss of function phenotype is specific and that the *fall*-GFP fosmid is functional. The *fall* phenotype can also be rescued by re-expressing *fall*-HA only in the developing IFMs by flight muscle specific *Act88F-GAL4* and *UAS-fall-HA* (Figure 1N-P). Together, these data demonstrate that *fall* is required specifically in flight muscles to prevent flight muscle atrophy.

Fall is required for sarcomere growth during development

In order to identify the cause for the muscle loss in *fall* mutants, we looked at IFM development during pupal stages. We found that flight muscle fibers look normal at 48 h APF in *fall* mutants and also the mutant sarcomeres have a wild-type morphology reaching a length of about 2.0 μm at 48 h APF (Figure 2A, E, I, M, R, W, Ab). This shows that myofibrillogenesis initiates normally and early sarcomeres are assembled. Also flight muscle innervation appears normal, ruling out that the atrophy is caused by missing innervation (Supplementary Figure 2). However, from 56 h APF onwards a progressive detachment of the IFMs in *fall*[1] was observed in most cases from the

posterior attachment sites (Figure 2B-D, F-H) and a severe muscle thinning resulting in muscle rupturing in *fall[2]* mutant IFMs (Figure 2I-L). Interestingly, this is accompanied by a severe sarcomere phenotype. At 48 h APF, most of the about 300 immature sarcomeres in each myofibril have been built in wild-type and *fall* mutant IFMs and are 2.0 μm long. Whereas in wild type the sarcomeres grow synchronously after 48 h APF to reach a length of about 3.0 μm by 72 h APF, sarcomere length growth is entirely blocked in *fall* mutants remaining at about 2.0 μm in length (Figure 2M-Ab). This is accompanied by actin accumulations found in *fall[2]* myofibrils from 56 h onwards (Figure 2X-Z). This demonstrates a severe sarcomere maturation defect in *fall* mutants. As the pupal thorax and the attached muscle fibers also grow in length during this sarcomere maturation phase, the *fall[1]* muscles detach from the attachments sites and the *fall[2]* mutants fibers rupture within the fiber, as they cannot support the required growth. Re-expression of *fall* with the late flight muscle specific driver *fln-GAL4*, which is only expressed after 48 h cannot rescue this fiber rupture phenotype (Supplementary Figure 1). Together, these data suggest that Fall is required in the flight muscles to support sarcomere maturation and to prevent muscle fiber rupturing from 48 h APF onwards.

***fall* mutant muscles produce uncontrolled high forces**

The muscle tearing and detachment phenotypes suggest an abnormally high force being present in the developing *fall* mutant flight muscles. In order to test if the phenotype is indeed dependent on force produced by the developing sarcomere, we crossed the *fall* mutants into a *Mhc[10]* background, which lacks most of the force producing Mhc in the flight muscles (Cripps et al., 1999). As expected, the muscle atrophy phenotype is entirely rescued in *Mhc[10];fall* double mutant flies (Figure 3A-F), suggesting that fiber tearing and detachment are indeed caused by uncontrolled high forces.

Muscle attachments need to be able to withstand high forces and integrins are central players for muscle-tendon attachments (Brown, 2000). In order to rule out that attachment sites are abnormally weak in *fall[1]* mutants, we stained for integrin complex components, however did not find any difference to wild type at 48 h APF, thus not explaining the detachment at 56 h (data not shown). We further tried to artificially strengthen the attachment sites by over-expression of a dominant active form of Rap1 or providing an additional genomic copy of the IPP complex member Pinch, both of which were shown to strengthen muscle attachment sites in flies (Ellis et al., 2013; Pronovost et al., 2013; Vakaloglou et al., 2016). However, both strategies did not rescue the muscle

detachment phenotype of *fall[1]* mutants (Supplementary Figure 2D-H), suggesting that a weakness of the attachment sites is not the cause for the detachment phenotype.

In order to investigate the detachment phenotype closer, we labeled the tendon cell membrane by expressing CD8-GFP with the tendon specific *stripe-GAL4*. At 56 h APF, the wild-type tendon cells show straight extensions from their basal side to the ends of the muscle fibers, where the integrin containing attachment sites are located (Figure 3G,I,J). Surprisingly, tendon cell membrane labeled by the CD8-GFP remains at the detached end of the *fall[1]* mutant muscle fibers at 56 h APF, showing that the tendon extensions rupture into pieces and the connection to the tendon cell body is lost. This is the case for both anteriorly occurring detachments, which are more rare, and posteriorly occurring detachments, which are frequently observed (Figure 3H,K-N). Thus, the muscle-tendon junction remains intact and integrin remains at the end of the muscle fibers (Figure 3K"-L"). The remaining tendon cell debris attached the *fall* mutant muscle fiber ends are degraded until 72 h APF (Supplementary Figure 3). As this phenotype can be rescued by muscle specific expression of *fall* (see Figure 1) these data strongly suggest that the forces produced in *fall[1]* flight muscles are too high to be counteracted by normal tendon cells, resulting in tendon cell rupturing and eventually muscle atrophy.

Fall is expressed in body muscle nuclei

Having seen this dramatic developmental defect in *fall* mutant flight muscles we wondered where does Fall protein localize. To address this, we first analyzed Fall expression using the functional Fall-GFP fosmid construct. We found that Fall-GFP is expressed in the nuclei of all adult body muscles, including both indirect flight muscle types, the dorso-longitudinal (DLM) and the dorso-ventral muscles (DVM), as well as jump and leg muscles. Fall-GFP is however not expressed in visceral muscles or any other cell types (Figure 4A-E). To address the temporal dynamics of Fall expression, we investigated Fall localisation during flight muscle development. We found that Fall-GFP is not yet expressed during the initiation of myofibrillogenesis at 30 h APF (Figure 4F). Fall-GFP expression can be detected at low levels at 34 h in the flight muscle nuclei, its expression strongly increases until 48 h and remains high until 72 h APF (Figure 4G-J). These localization studies demonstrate the Fall-GFP is indeed a body muscle specific nuclear protein, whose expression is strongly induced during the myofibril and sarcomere maturation phase, nicely correlating with the observed phenotypic defect in sarcomere maturation in *fall* mutants.

Fall requires its BTB and Zn-finger domains to function

After we had established the nuclear localization of Fall during muscle development we wanted to investigate the role of the different domains of Fall for sarcomere maturation. To address this important point, we used the CRISPR generated *fall*[2] allele that in addition to the STOP cassette and the 3x-P3dsRed marker also contains two flanking attP sites, which can be used for phiC31 recombinase mediated cassette exchange (Venken et al., 2011). We used this strategy to either insert a *fall* cDNA, tagged at the C-terminus with HA and fused in frame with the remaining 38 amino acids of the endogenous first *fall* exon, or a *fall* cDNA that is lacking the BTB or both zinc finger domains (Figure 5A). All three HA tagged proteins are expressed at a comparable level and the flies are homozygous viable. However, only the *fall*-CDS-HA flies can fly, whereas *fall*- Δ BTB-HA and *fall*- Δ ZFs-HA are entirely flightless (data not shown). We analyzed the flight muscle morphology of these three different *fall* alleles and found that *fall*-CDS-HA has an entirely normal IFM morphology at 48 h and 72 h APF and also its sarcomeres grow normally to a length of about 2.8 μ m at 72 h APF (Figure 5B,E,H,K,N). In contrast, both *fall*- Δ BTB-HA and *fall*- Δ ZFs-HA alleles display the typical myofiber rupture phenotype at 72 h APF likely caused by severe sarcomere growth defects with the sarcomere length remaining close to 2 μ m at 72 h APF (Figure 5B-N). This demonstrates that Fall protein needs both, its BTB and its zinc finger domains, to support sarcomere maturation and flight muscle fiber integrity during muscle development.

The BTB domain of BTB zinc finger proteins often mediates protein-protein interactions, whereas its zinc fingers are often implicated in binding to DNA (Siggs and Beutler, 2012). We investigated the distribution of the full length and the deletion proteins during flight muscle development and interestingly found that in contrast to Fall-HA, Fall- Δ BTB-HA is defective in nuclear enrichment. It apparently can enter the nuclei, but is not retained there, resulting in an equal distribution between cytoplasm and nuclei (Figure 5O,P). On the other hand Fall- Δ ZFs-HA is normally localized in the muscle nuclei (Figure 5Q) but non functional, consistent with the interpretation that its zinc fingers are not required for nuclear localization but rather for its function, possibly for its interaction with DNA.

Fall is a transcriptional regulator at sarcomere maturation stages

We found a sarcomere maturation defect in *fall* mutants. In order to directly test if Fall is indeed a transcriptional regulator during the stages of sarcomere maturation we dissected indirect flight muscles from wild type and *fall*[1] mutants at 30 h and 48 h APF and

performed mRNA-Seq analysis. In accordance with the wild type morphology and the absence of Fall expression at 30 h APF, we find very few transcriptional changes at 30 h APF (Figure 6A, Supplementary Table 1). In contrast, the expression of a large number of genes are changed at 48 h APF, including a large number of sarcomeric components being down-regulated in *fall[1]* (Figure 6B, Supplementary Table1). This demonstrates that Fall is indeed a transcriptional regulator during sarcomere maturation stages.

We performed a bioinformatic analysis using GO categories and additional informative gene function lists. We found a significant down-regulation of transcription factors and sarcomeric components, as well as genes identified as the 'salm-core' set of genes dependent on Spalt function in flight muscle (Spletter et al., 2015). Additionally, we found a significant change in the mitochondrial metabolism, with glycolysis components being up-regulated, as well as pyruvate and fat metabolism being down-regulated. One of the most striking differences is found by the analysis of the 'gene cluster 22 set', which was defined as a set of genes being strongly induced during sarcomere maturation stages containing a strong enrichment of sarcomeric as well as mitochondrial components (Figure 6C) (Spletter, Zhang & Schnorrer, in prep.). The fly GO-term memberships are somewhat ill-defined, with many bona-fide sarcomeric proteins not listed with a sarcomeric GO-term, but rather with cytoskeleton related GO-terms. We find many of them including *ZASP66*, *starvin*, *sarcomere length sort (sals)*, the titin homolog *sallimus (sls)* and *bent (bt)*, the M-line components *obscurin (unc-89)* and *mask*, as well as *paxillin (pax)* being down-regulated in *fall[1]* muscles at 48 h APF (Figure 6D).

It is challenging to verify all these mRNA expression changes at a protein level, as total proteome detection with mass-spectrometry from muscles has still limitations to quantify all proteins. Nevertheless, we have been able to confirm the down-regulation of a significant number of proteins, including the sarcomeric components Strn-Mlck and TpnC4 at 72 h APF, whereas other sarcomeric components, including TpnC73F and TpnC41C are aberrantly up-regulated (Figure 6E, Supplementary Table 2). We used a genomic GFP fusion in *ZASP66* to confirm its down-regulation at sarcomeric Z-discs of 72 h APF flight muscles (Figure 6F). We had antibodies working for western blot detection available for Flightin (Fln) and Gelsolin (Gel) and could confirm its down- and up-regulations, respectively (Figure 6G). Taken together, these data strongly suggest that Fall is essential to regulate the induction of a large number of sarcomeric components during the stages of sarcomere growth in order to enable proper sarcomere maturation.

Fall localises to nuclear speckles

By inspecting the Fall-GFP or Fall-HA nuclear localizations in greater detail, we realized that Fall localizes non homogenously throughout the muscle nuclei. These patterns reminded us of stress induced nuclear granules, which have been reported to regulated chromatin landscapes (Gómez-Díaz and Corces, 2014). In order to test if the Fall localization pattern in the nuclei is dynamic we treated semi-dissected hemi-thoraces of 48 h APF for 30 min with PBS or PBS plus 250 mM NaCl and compared them to directly fixed samples. Interestingly, we find that the small Fall-GFP speckles that are found when directly fixed or treated just with PBS dynamically fuse to large speckles which preferentially cluster at the nuclear periphery after incubation in high salt (Figure 7A-E).

Generally, the recruitment of wild-type Fall protein into large speckles upon high salt treatment is not dependent on the developmental stage of the muscle and can be induced at 48 h, 72 h or 1 day adults with similar efficiencies (Figure 7C,L-O). It is also not only caused by high salt treatment but also by 30 min incubation with 500 mM sucrose (Supplementary Figure 4A), suggesting that large Fall speckles are induced by different kinds of osmotic stress. The recruitment of Fall into speckles is not a general phenomenon of all nuclear proteins as for example the localization of the heterochromatin protein HP1 does not change upon osmotic stress (Supplementary Figure 4B,C). These findings are in accordance with Fall being a dynamic transcriptional regulator, whose activity may be adjusted according to the status of the cell.

In order to identify the domains in Fall required for the recruitment into large nuclear speckles, we treated 72 h APF thoraces of *fall*-CDS-HA and the *fall* deletions alleles with high salt. We found that Fall-CDS-HA and Fall- Δ -ZFs-HA are recruited into large speckles at 72 h APF, but the small amount of Fall- Δ BTB-HA that is present in the nucleus at 72 h APF is not localized to any speckled pattern (Figure 7F-H). This again demonstrates the importance of the Fall BTB domain for the normal localization of Fall, whereas the zinc finger domains are functionally required but do not show a difference in the localization pattern.

Fall localization is distinct from insulator bodies defining a new nuclear body

A similar dynamic re-localization upon osmotic stress as we found for Fall was described for a class of proteins that are generally coined insulator bodies, including CP190, BEAF-32, Su(Hw) and Mod(mdg4)67.2 (Schoborg et al., 2013). Interestingly, CP190 and Mod(mdg4)67.2 also contain BTB domains. As these components are not expressed

in a tissue specific manner but rather ubiquitous, we hypothesized that Fall may be a muscle specific component of the insulator complex. We used the Fall-GFP or Fall-HA flies to investigate the recruitment of Fall into insulator bodies defined by BEAF32, CP190, Su(Hw) and Mod(mdg4)67.2. Surprisingly, we found that although all five components are recruited to large speckles upon osmotic stress, the Fall speckles or droplets are in close proximity to neither overlap with the CP190 or BEAF32 pattern (Figure 8A-D) nor with the Su(Hw) or Mod(mdg4)67.2 speckles (Supplementary Figure 5A-D). We also observed that the overlap between the bona-fide insulator complex members CP190 and BEAF-32 is not complete upon insulator or droplet formation, whereas CP190 and Mod(mdg4)67.2 overlap very well in flight muscle after osmotic stress (Figure 8E,F). These data may suggest that there are a larger number of nuclear speckles or insulator bodies existing. Furthermore, they may share important features with the recently described nuclear lipid droplets, which supposedly are used to segregate different nuclear proteins (Mitrea and Kriwacki, 2016; Zhu and Brangwynne, 2015). It is important to note that most of the generally accepted important insulator complex components appear to be dispensable for normal development. Null or strong hypomorphic alleles for *CTCF*, a additional core insulator member (Mohan et al., 2007), *BEAF-32* or *CP190* are viable or only pharate lethal and their flight muscles do not display any obvious gross defects (Supplementary Figure 5E-H). Together, these data suggest that Fall, in contrast to the classical insulator complex components, has an essential role for sarcomere maturation during flight muscle development and may fulfill this role by using a novel type of nuclear speckle or droplet domain to function.

Discussion

In this study we have investigated the mechanism controlling the transcriptional induction of sarcomeric components to enable normal sarcomere maturation and growth during flight muscle development of *Drosophila*. We have identified the BTB-zinc finger protein Fallen Angel as essential transcriptional regulator, which is crucial for the boosted expression of a subset of sarcomeric components after immature myofibrils and sarcomeres have been built. This ensures that sarcomeres can mature to its normal pseudo-crystalline pattern found in the adult.

In a recent systematic developmental mRNA sequencing study we found a biphasic expression for many sarcomeric components, with a lower induction before the initiation of myofibrillogenesis resulting in the formation of immature sarcomeres. This is followed by a substantial transcriptional boost during sarcomere maturation, coinciding with sarcomeric growth from 1.7 μm to 3.2 μm in the adult (Spletter, Zhang & Schnorrer, in prep). This work defined a cluster containing a subset of sarcomeric components and a number of mitochondrial components to be strongly induced during sarcomere maturation. Interestingly, this cluster of genes strongly depends on Fall protein function, with many sarcomeric and mitochondrial components not being induced in *fall* mutants. This defined Fall as a novel transcriptional regulator with particular phenotype, distinct from all other known transcriptional regulators in muscle, a specific regulator of sarcomere maturation.

Fall requires both its BTB domain and the zinc finger domains to fulfill its function. As the BTB domain is required for the correct nuclear localization of Fall and is a likely interaction platform with other proteins (Siggs and Beutler, 2012), it is fair to assume that Fall does not act alone but within a larger molecular complex to regulate its targets. To date, we can only speculate about the nature of this complex. We have shown that Fall dynamically localizes to a novel class of nuclear bodies, which are in proximity but distinct from the well described insulator bodies, that similarly to Fall respond to osmotic stress (Schoborg et al., 2013).

Our findings are consistent with a recent model of liquid phase separation controlling transcriptional decisions in general and the formation of so called super-enhancers in particular (Hnisz et al., 2017). At super-enhancers, a large amount of transcriptional activators are concentrated in a cooperative manner, potentially using principles of liquid phase separation. These enhancers are particularly active resulting in very high

expression rates (Lovén et al., 2013). As sarcomeric proteins require very high induction rates in order to enable normal sarcomere maturation a similar liquid phase complex containing Fall and other factors may boost the expression of these components during sarcomere maturation.

We have shown here that the correct stoichiometry of the sarcomeric components is critical for normal muscle force production during development. In the likely hypomorphic *fall[1]* allele sarcomere growth is entirely blocked, but sarcomeric architecture is initially relatively preserved, no actin accumulations are visible. These *fall[1]* mutant fibers can produce such high forces from 56 h APF onwards that the attached tendon cells rupture into pieces, showing that in wild type muscle force production must be tightly regulated. The muscle fibers of the *fall* null allele *fall[2]* rupture within the muscles, likely because the compromised sarcomeric and myofibril architecture with the actin accumulations cannot withstand the high forces, thus the muscle and not the tendon cells rupture. The result in both alleles is the same: the muscle fibers are disconnected from the tendons cells and undergo severe muscle atrophy. This demonstrates the importance of controlled sarcomere maturation and force production for muscle function and muscle cell survival.

To date, it is unclear if vertebrates have a functional ortholog of Fall. BTB-zinc fingers are a large and diverse family, in which only the BTB and the zinc-finger domains are preserved (Siggs and Beutler, 2012). With regular BLAST searches a Fall ortholog in vertebrates cannot be defined. However, a recent study identified mutations in human ZBTB42 as a cause for a lethal congenital contracture syndrome, caused by uncontrolled muscle hyper-contraction (Patel et al., 2014). Knock-down of the zebrafish ZBTB42 also resulted in a severe muscle phenotype (Patel et al., 2014). Thus, we speculate that ZBTB42 might be a functional ortholog of Fall in human skeletal muscle, ensuring the proper stoichiometry of the sarcomeric components to control force production during skeletal muscle development.

Acknowledgement

We thank the fly community for sharing fly stains and antibodies. We are grateful to Reinhard Fässler for generous support and to Bettina Stender and Nicole Plewka for their excellent technical assistance. This work was supported by the Max Planck Society, the European Research Council under the European Union's Seventh Framework Program (FP/2007-2013)/ERC Grant 310939, a Career Development Award from the Human Frontier Science Program (F.S.), the EMBO Young Investigator Program (F.S.).

Figures

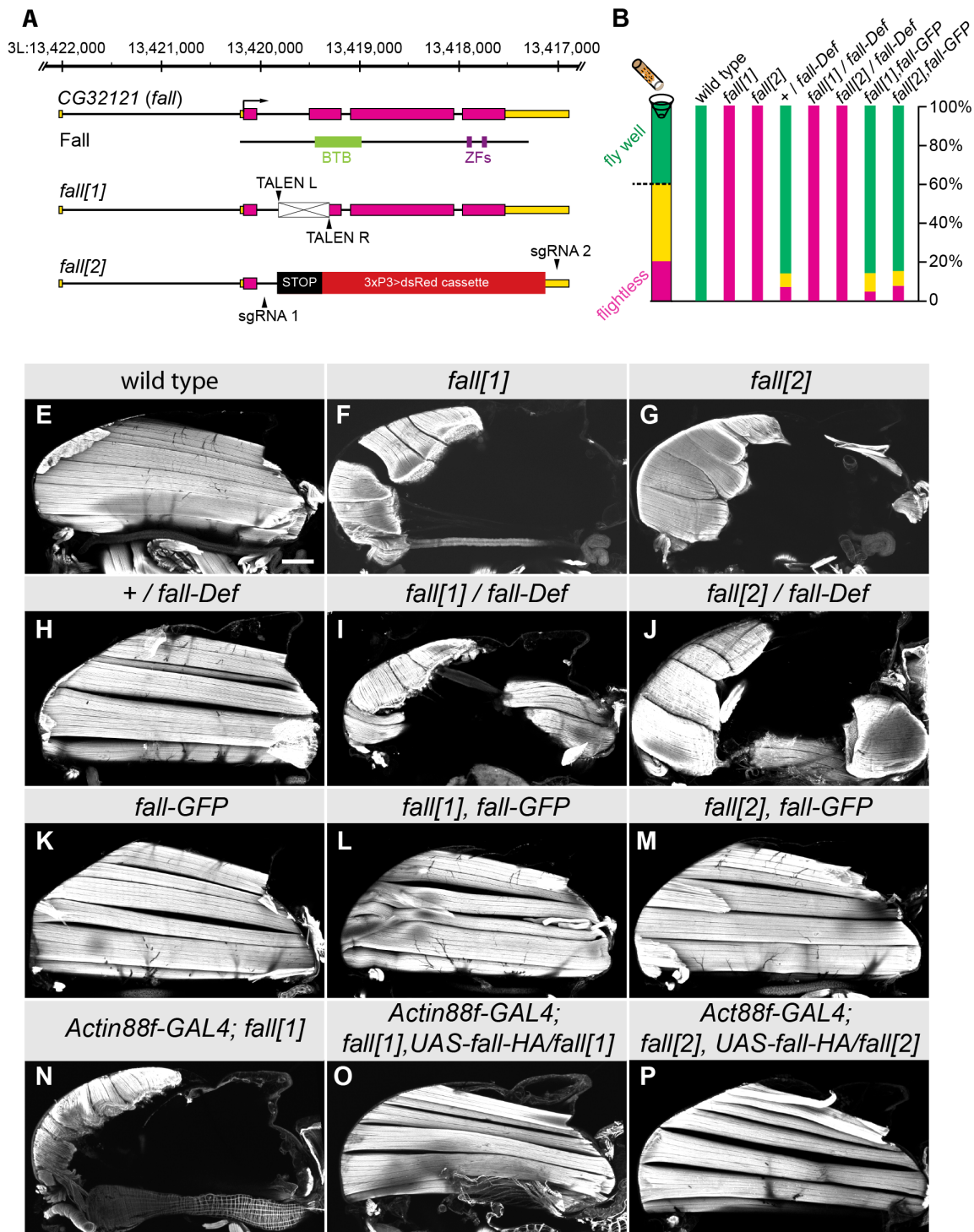


Figure 1: Fall is required for IFM development or *fall* is required in muscle to prevent muscle atrophy

A: Gene and protein structure of fall;

B: TALEN mediated deletion scheme of *fall*[1];

C: CRISPR based deletion scheme of *fall*[2];

Appendices

D: Flight test;

E-J: Dissected adult hemithoraces of wild type (E), *fall[1]* (F), *fall[2]* (G), wild type over *fall* deficiency (H), *fall[1]* over *fall* deficiency (I), *fall[2]* over *fall* deficiency (J);

K-M: Rescue of *fall* mutation with a GFP tagged Fosmid line Fall-GFP: *fall-GFP* (K), *fall[1]*, *fall-GFP* (L), *fall[2]*, *fall-GFP* (M);

N-P: Tissue specific rescue of *fall* alleles with IFM specific driver: *Actin88f-GAL4*; *fall[1]* (N), *Actin88f-GAL4*; *fall[1]*, *UAS-fall-HA/fall[1]* (O), *Actin88f-GAL4*; *fall[2]*, *UAS-fall-HA/fall[2]* (P).

For flight test, around 20-30 flies were flipped into the 1 meter long cylinder and the landing position was recorded. The respective percentages of each zone were calculated as the number of flies in that zone over total tested flies. Scale bar represents 100 μ m.

Appendices

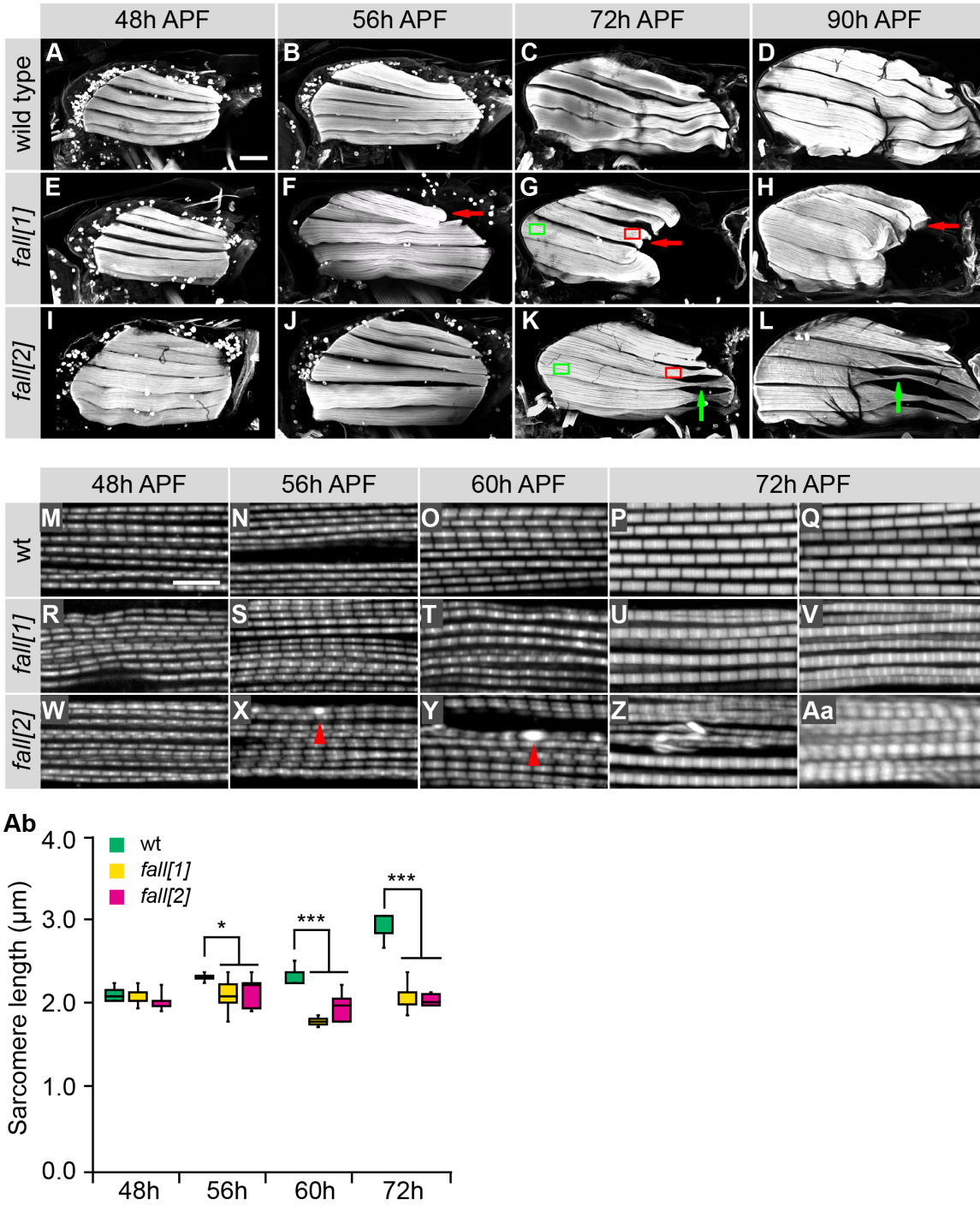


Figure 2: *fall* is required for sarcomere maturation during muscle growth

A-D'': Rhodamine-phalloidin staining of developing IFM from 48h APF to 90h APF in wild type (A-D), *fall[1]* (A'-D'), *fall[2]* (A''-D''); The red arrows point to contracted muscle fiber tips; green arrows point to ruptured muscle fibers.

E-I'': Developing myofibrils and sarcomeres from 48h APF to 90h APF in wt (E-I), *fall[1]* (E'-I'), *fall[2]* (E''-I''); Actin blobs are frequently seen at 56h APF in *fall[2]* (white arrowhead). H' and H'' were taken from green-boxed region in C' and C'', respectively. I' and I'' were taken from red-boxed region in C' and C''. Scale bar

Appendices

represents 100 μm in A-D'' and 5 μm in E-I''. Images in A-D'' and E-I'' are cropped to the same size, respectively.

J: Sarcomere length quantification, * $P < 0.05$, *** $P < 0.001$, two tailed unpaired Student's t-test. Error bar represents the \pm s.d. from at least three samples for each group.

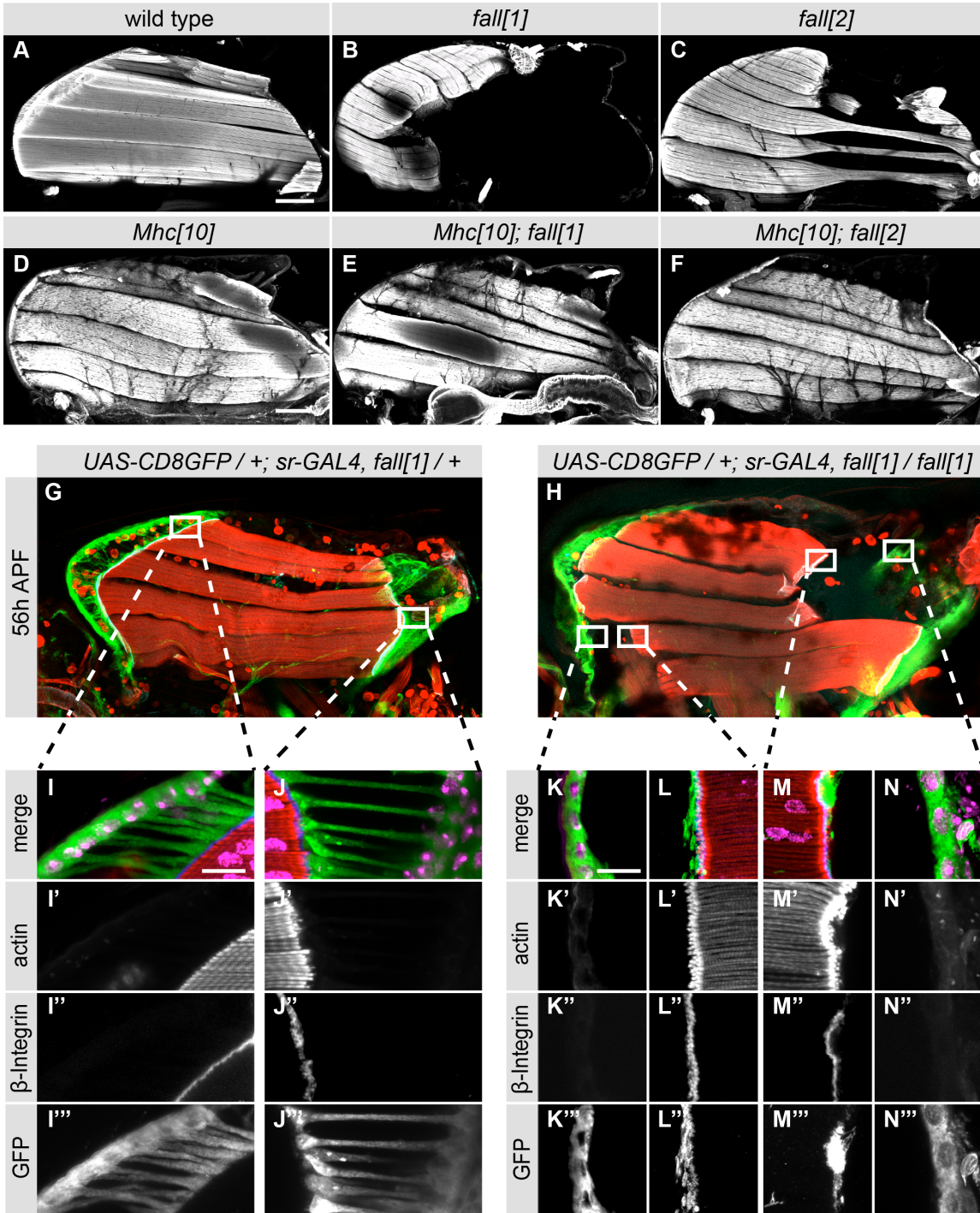


Figure 3: *fall* loss of function causes muscle hypercontraction during sarcomere maturation

A-F: Dissected young adult hemithoraces of wild type (A), *fall[1]* (B) and *fall[2]* (C) *Mhc[10]* (D), *Mhc[10]; fall[1]* (E), and *Mhc[10]; fall[2]* (F). Note that intact IFM fibers are retained in E-F.

G-N''': Tendon cell and extensions are labelled with *UAS-CD8-GFP; stripe-GAL4*. Dissected 56h APF pupal hemithoraces were stained for GFP, β-PS, and shot of *UAS-*

Appendices

CD8-GFP; stripe-GAL4, fall[1]/+ (G), *UAS-CD8-GFP; stripe-GAL4, fall[1]/+* (H). High magnification of muscle tendon junction (white boxed region) at anterior end (I-I''') and posterior end (J-J''') of *UAS-CD8-GFP; stripe-GAL4, fall[1]/+*; High magnification of anterior tendon cells (K-K''') and ruptured muscle tips (L-L'''), as well as posterior ruptured muscle tips (L-L''') and tendon cells (K-K''') of *UAS-CD8-GFP; stripe-GAL4, fall[1]/fall[1]*. Note that β -PS specifically localizes to muscle fiber tips and tendon cell extension debris was seen at both anterior and posterior ruptured muscle tips. Scale bars are 100 μ m in A-F and 10 μ m in I-J''' and K-N'''. Images of A-C, D-F, I-J''', and K-N''' are cropped to the same size, respectively.

Appendices

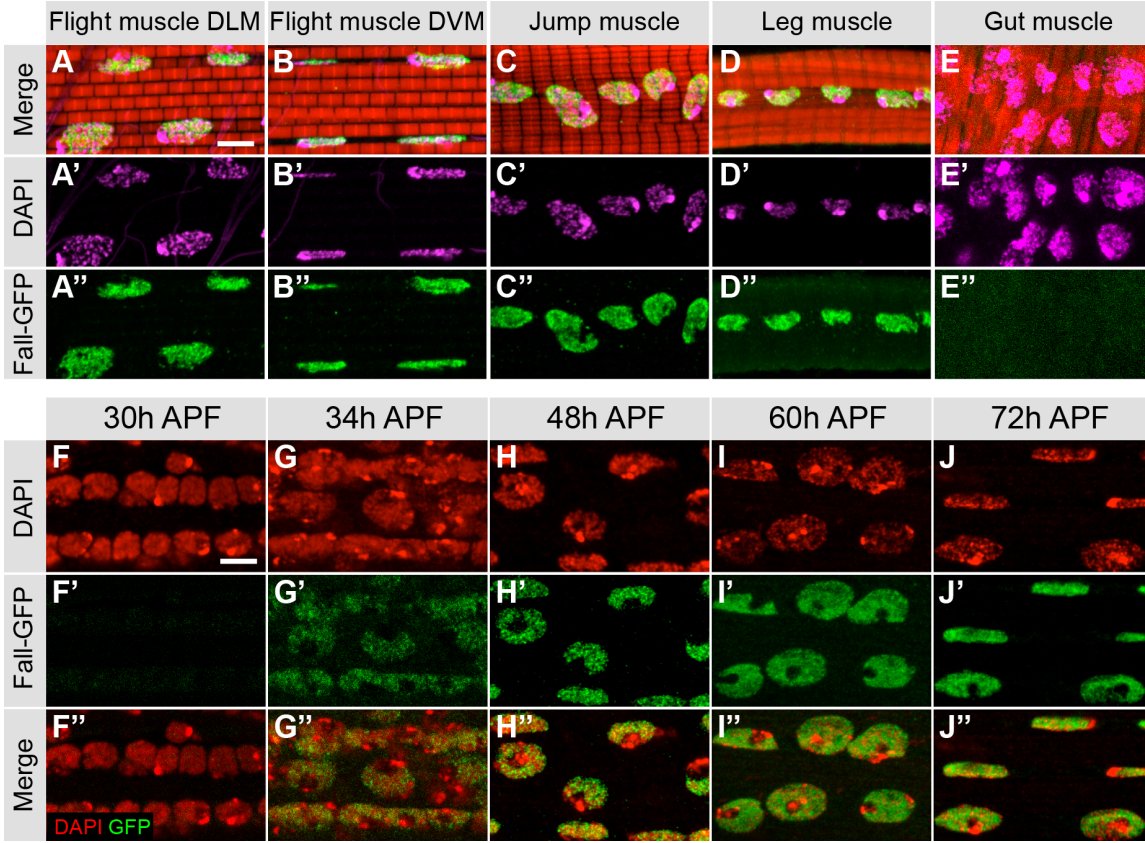


Figure 4: Fall is localised in body muscle nuclei

A-E'': Dissected *fall-GFP* adult body muscles are stained for GFP and DAPI. GFP are seen in DLM (A-A''), DVM (B-B''), jump muscle (C-C''), and leg muscle (D-D'') but absent in visceral gut muscle (E-E'').

F-J'': Dissected *fall-GFP* developing IFMs are stained for GFP and DAPI at 30h APF (F-F''), 34h APF (G-G''), 48h APF (H-H''), 60h APF (I-I''), and 72h APF (J-J''). Fall-GFP starts to express at 34h APF and reach high level at 48h APF in IFM nuclei. Note that Fall-GFP displays a discrete small dot pattern.

Scale bars are 5 μm in all cases. Images of A-E'' and F-J'' are cropped to the same size, respectively.

Appendices

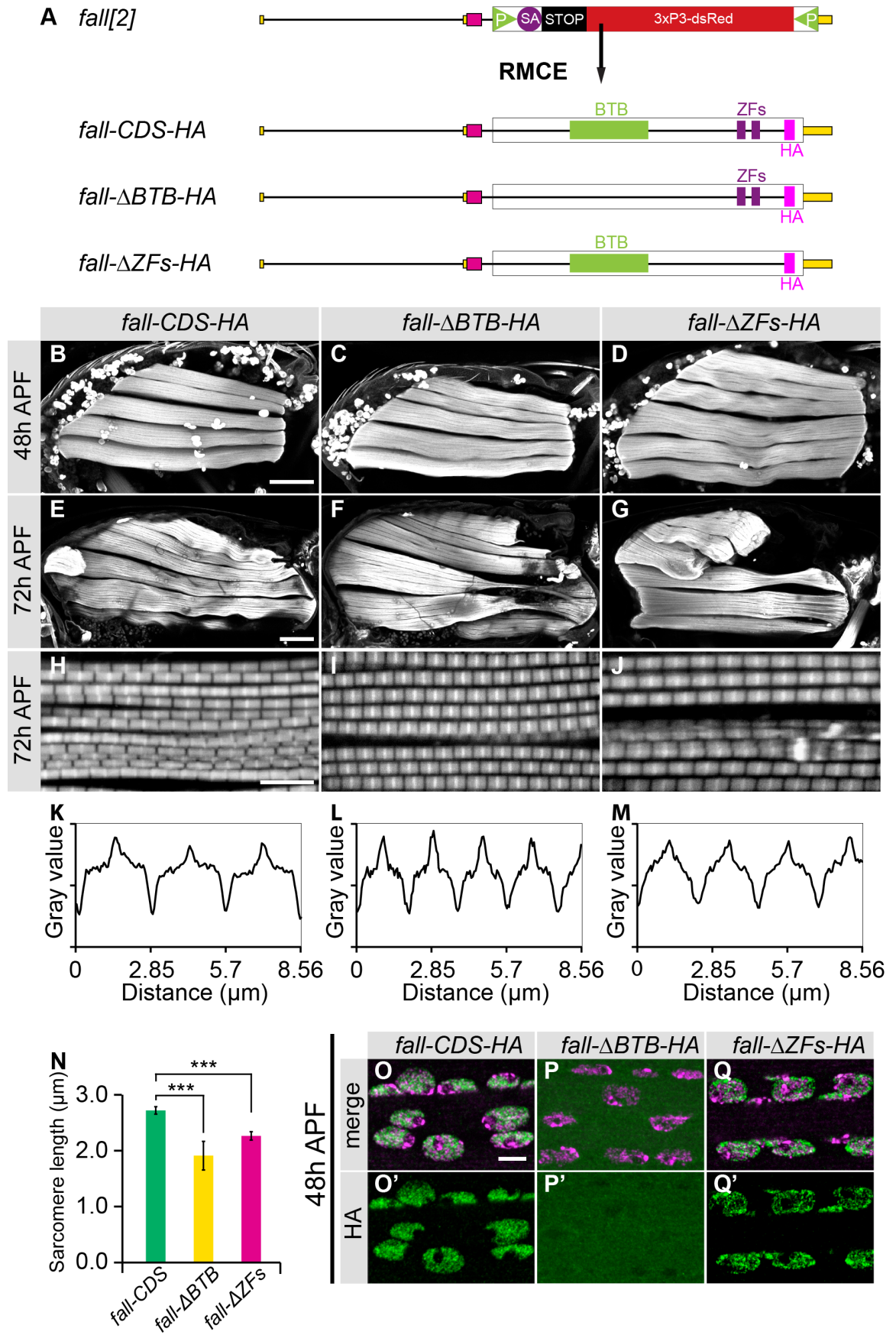


Figure 5: Both BTB and ZFs are required for Fall function

Appendices

A: *fall-CDS-HA* (B), *fall-ΔBTB-HA* (C), and *fall-ΔZFs-HA* (D) are insert to *fall* endogenous locus by phiC31 integrase mediated RMCE.

B-G: Dissected 48h APF and 72h APF pupal hemithoraces of *fall-CDS-HA* (B, E), *fall-ΔBTB-HA* (C, F), and *fall-ΔZFs-HA* (D, G), Muscle fibers ruptured in F and G.

H-J: Myofibrils and sarcomeres of *fall-CDS-HA* (H), *fall-ΔBTB-HA* (I), and *fall-ΔZFs-HA* (J) at 72h APF.

K-M: Relative intensity profile of a representative myofibrils from image H (K), image I (L), and image J (M). Sarcomeres are shorter in I/L and J/M.

N: Quantification of sarcomere length of *fall-CDS-HA*, *fall-ΔBTB-HA*, and *fall-ΔZFs-HA* at 72h APF. ***P < 0.001, two-tailed unpaired Student's t-test. Error bar represents the ± s.d. from at least three samples for each group.

O-Q': Localization pattern of Fall-CDS-HA (O-O'), Fall-ΔBTB-HA (P-P'), and Fall-ΔZFs-HA (Q-Q') at 48h APF. Note that Fall-ΔBTB-HA localizes to cytoplasm rather than nuclei. Scale bars represent 100 μm in B-D and 5 μm in H-J and O-Q'. Images of B-D, E-G, H-J, and O-Q' are cropped to the same size, respectively.

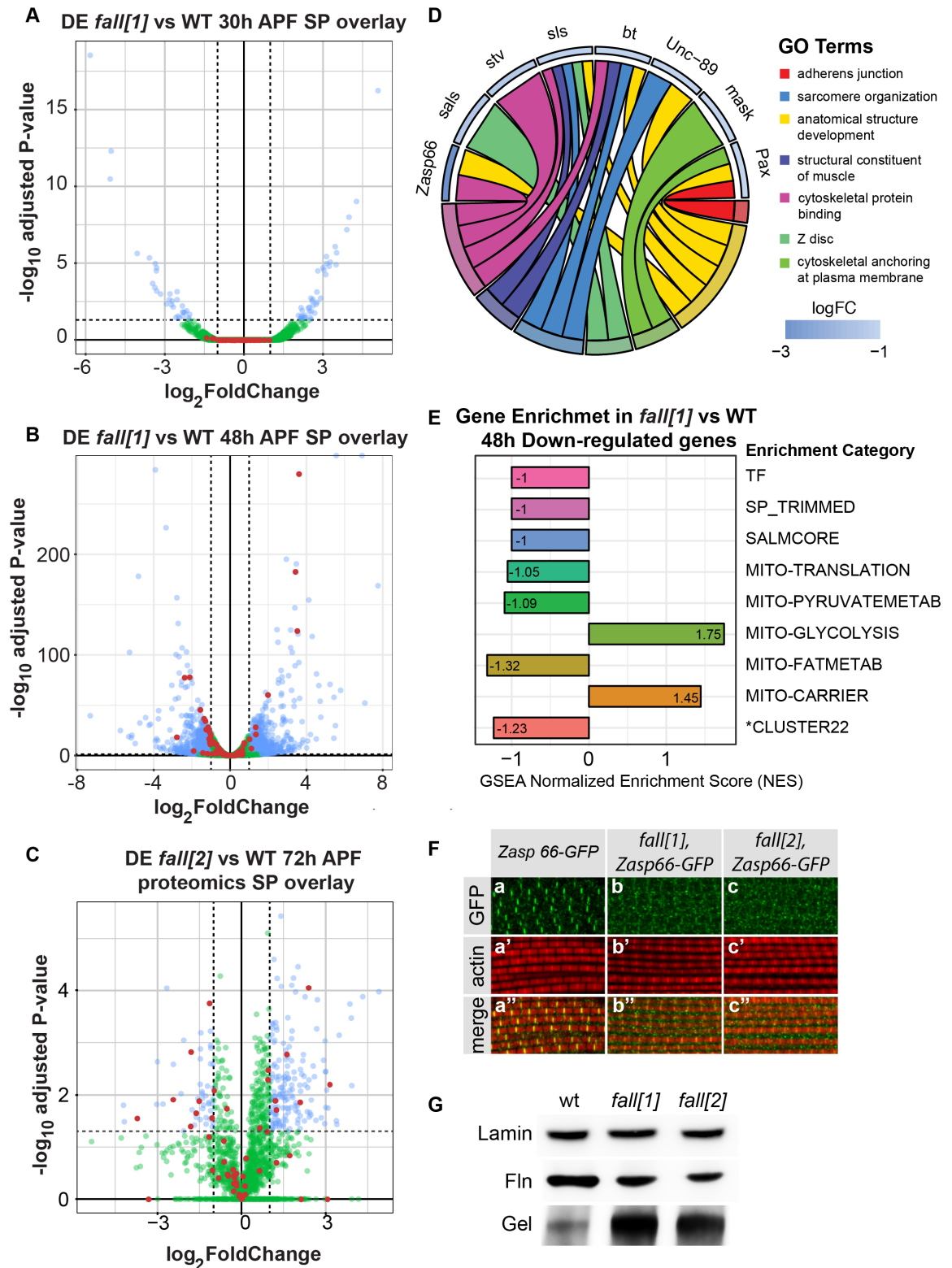


Figure 6: Fall regulates a large group of genes that are required for sarcomere maturation or Fall modulates transcription during sarcomere maturation stage.

A-B: Volcano plot of significantly changed genes (DESeq2, $p < 0.05$, $\log_2FC > 1$) of *fall[1]* vs *wild type* RNA-seq at 30h APF (A) and 48h APF (B). Blue dots indicate significant changed genes, green dots indicate insignificant changed genes, and red dots indicate

Appendices

overlay with sarcomere genes. Note, at 30h, only few genes are significantly changed and no sarcomere genes changed. At 48h, many genes changed and sarcomere genes tend to be low expressed in *fall[1]*.

C: Volcano plot of *fall[2]* vs *wild type* proteomics at 72h APF. Red dots indicate overlay with sarcomere genes.

D: GSEA analysis of user-specified gene sets including sarcomere genes, salmcore genes, cluster 22 (a genes set that are induced at sarcomere maturation stage, Maria unpublished) and mitochondria related genes. Note that the sarcomere genes, salmcore genes and cluster 22 genes are significant enriched to a down regulation trend in *fall[1]*. The trend of down regulation is indicated by Normalised Enrichment Score.

E: Significantly down regulated genes at 48h APF in *fall[1]* are analysed by GOElite. Representative significantly enriched GO terms and genes are picked and visualised by Chordplot.

F: Dissected 72h APF IFM from *ZASP66-GFP* (a-a''), *Fall[1]*, *Zasp-66* (b-b''), *fall[2]*, *Zasp66-GFP* (c-c''). In *fall[1]* and *fall[2]*, *Zasp66-GFP* level is significantly lower and no obvious periodic localization is seen.

G: Western blot of 56h APF *wild type*, *fall[1]*, *fall[2]* IFM protein samples, probed for Lamin (control), Flightin and Gelsolin. Flightin level in *fall[1]* and *fall[2]* are significantly lower than wild type, while it is opposite for Gelsolin.

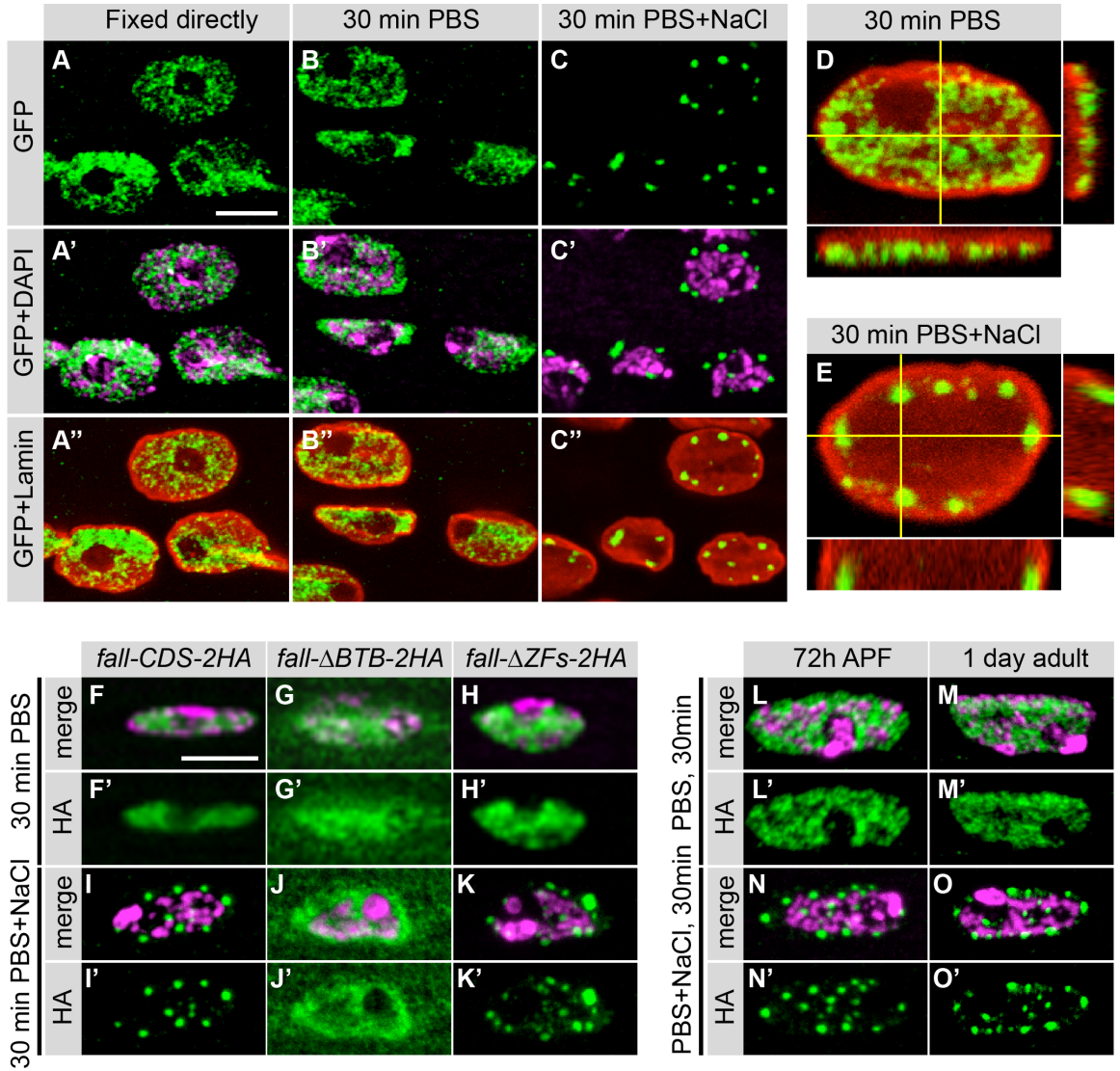


Figure 7: Fall forms nuclear body like structure or Fall localizes to liquid droplets in muscle nuclei upon osmotic stress

A-C'': Fall-GFP localization pattern at 48h APF without (A-B'') or with 30 min NaCl treatment (C-C''). Dissected IFM samples are stained for GFP, DAPI, and lamin.

D-E: Orthogonal view of single representative nucleus from untreated (D) and treated sample (E). Fall-GFP nuclear body localized to the nuclear membrane periphery that labeled by lamin.

F-H': Localization pattern of Fall-CDS-HA (F-F''), fall-ΔBTB-HA (G-G''), and fall-ΔZFs-HA (H-H'') at 72h APF without (F-H and F'-H') or with 30 min of 250 mM NaCl treatment (F''-H'' and F'''-H''). Fall-ΔBTB-HA failed to enter nuclear bodies after osmotic treatment.

I-J'': Fall-CDS-HA localization pattern at 72h APF (I-I'') and 1 day old adult (J-J''). Samples are stained for HA and DAPI.

Appendices

Scale bars represent 5 μm in all images. Images of A-C'' and F-J''' are cropped to the same size, respectively

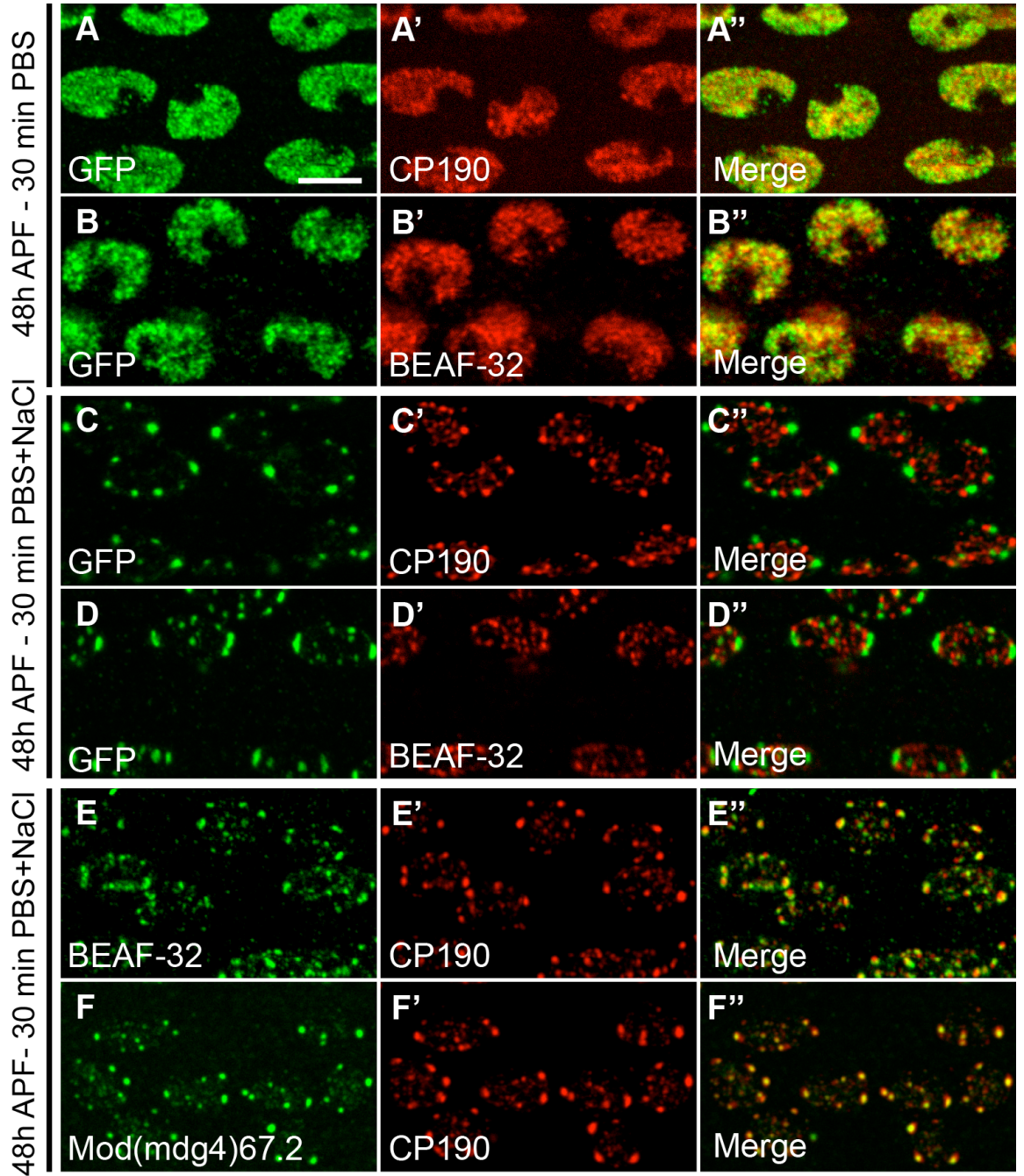


Figure 8: Fall-GFP nuclear body does not overlap with insulator bodies

A-B'': Dissected 48h APF IFMs from *fall-GPF* are stained for GFP (A and B) and CP190 (A'), and BEAF-32 (B').

C-D'': Dissected 48h APF IFMs from *fall-GPF* are treated with 250 mM NaCl for 30min and then stained for GFP (C and D) and CP190 (C'), and BEAF-32 (D'). In the merged image, Fall-GFP does not colocalize with CP190 or BEAF-32, but appears that associated with CP190 (C'').

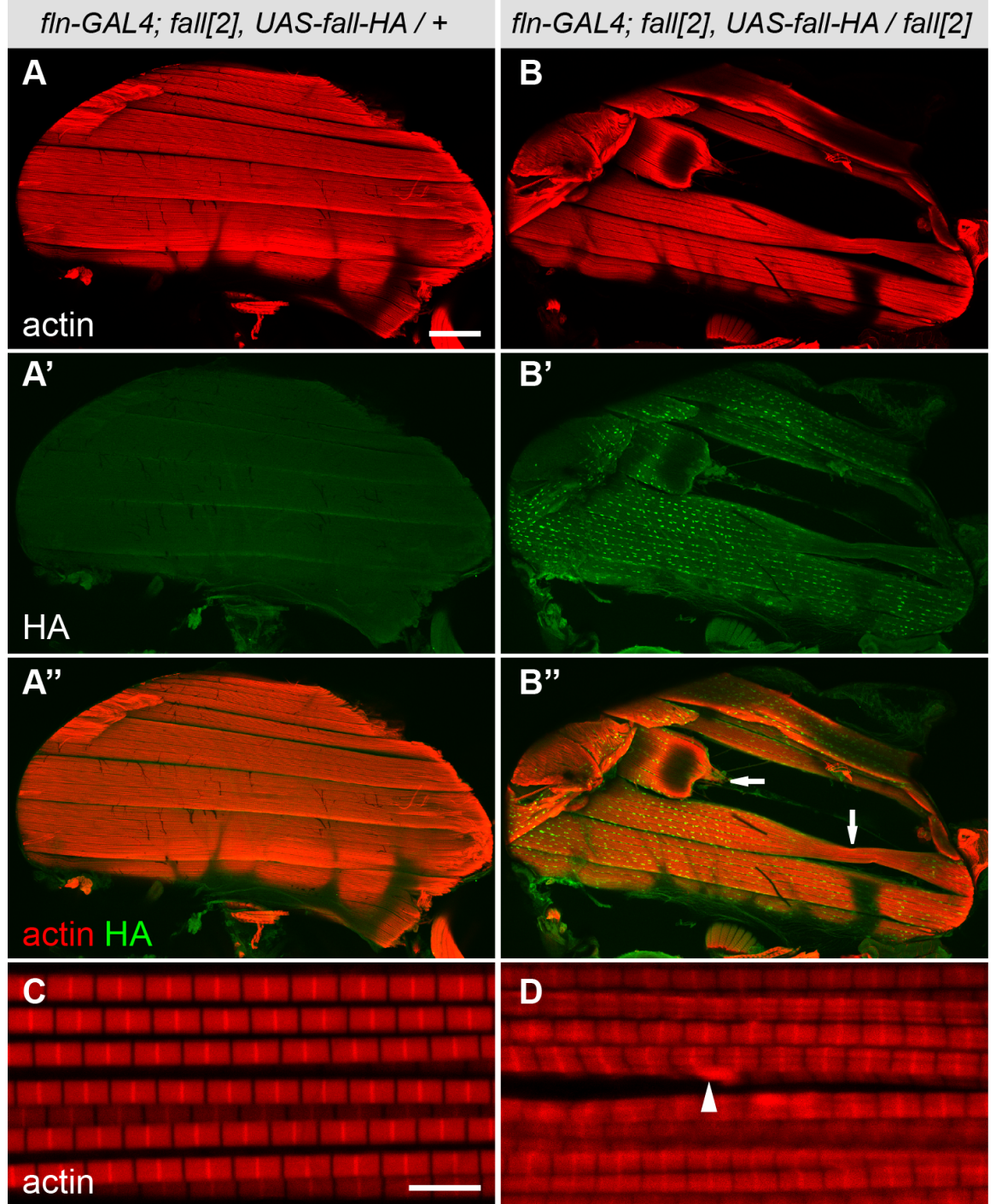
E-F'': Dissected 48h APF IFMs from wild type are treated with 250 mM NaCl for 30min and then stained for BEAF-32 (E), CP190 (E' and F'), and Mod(mdg4)67.2 (F). CP190

Appendices

and BEAF-32 (E'') are associated but not complete overlap while CP190 and Mod(mdg4)67.2 (F'') are completely overlapped.

Scale bars represent 5 μm in all images and all images are cropped to the same size

Supplementary Figures



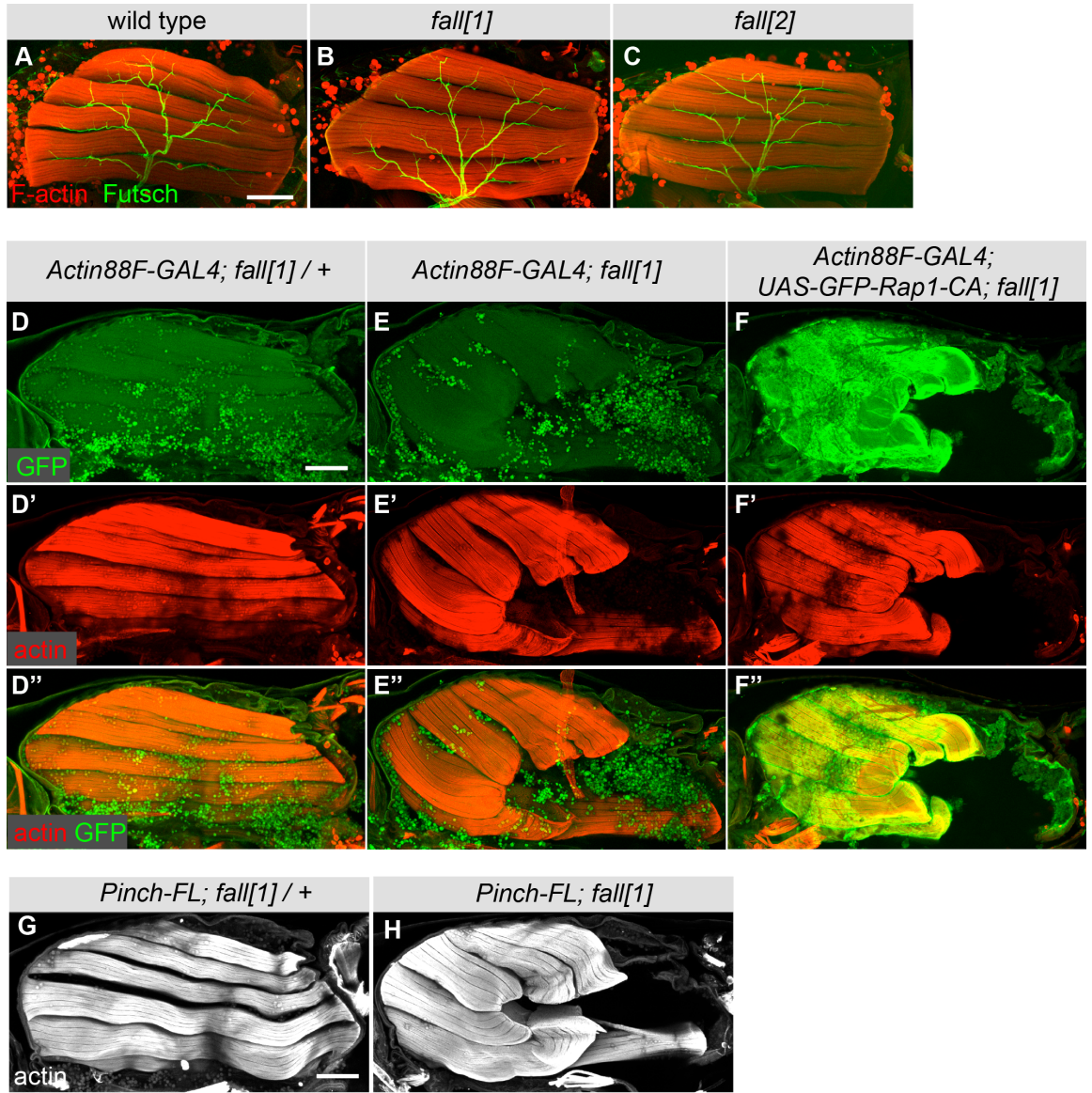
Supplementary Figure 1: *Fln-GAL4*; *UAS-fall-HA* failed to rescue *fall* mutants.

A-A'' and B-B'': Dissected adult hemithoraxes were stained for F-actin (A, B) and HA (A', B') of *fln-GAL4/+; fall[2], UAS-fall-HA/+* (A-A'') and *fln-GAL4/+; fall[2], UAS-fall-HA/fall[2]* (B-B'').

A''' and B''': Myofibril and sarcomere morphology of *fln-GAL4/+; fall[2], UAS-fall-HA/+* and *fln-GAL4/+; fall[2], UAS-fall-HA/fall[2]*.

Appendices

White arrows point to the ruptured muscle fibers and white arrowhead points to irregular sarcomeres. Scale bar is 100 μm in A-A'' and B-B'' as well as 5 μm in A'''-B'''.



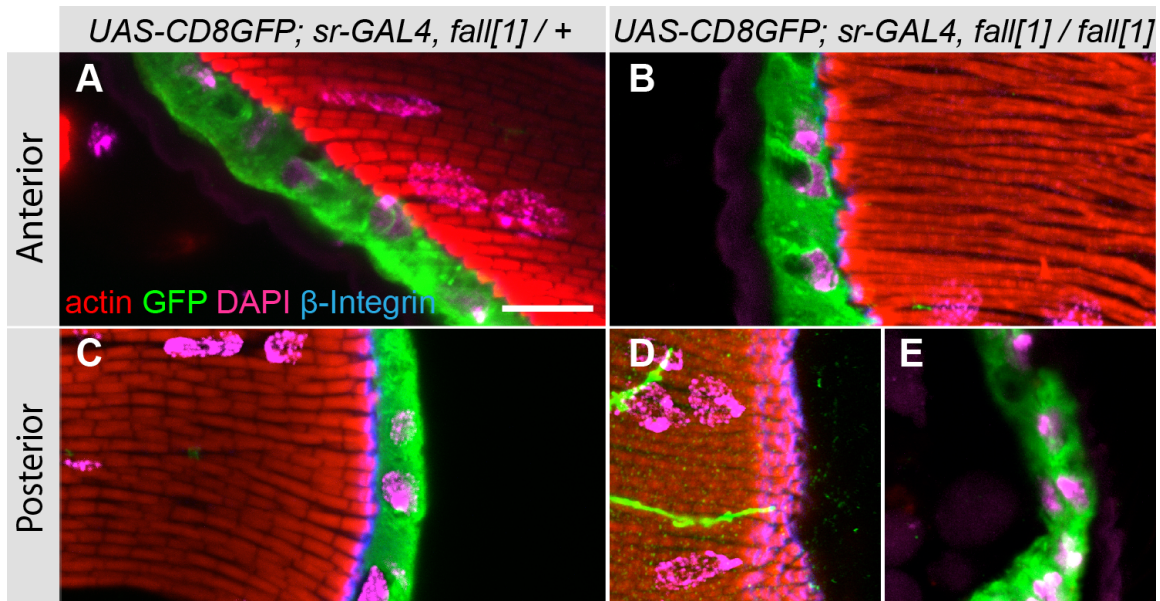
Supplementary Figure 2: IFM neuronal pattern in *fall[1]* and *fall[2]* is similar to wild type; MTJ reinforcement by overexpressing constitutively active form Rap1 or Pinch did not rescue muscle rupture in *fall[1]*

A-C: IFM neuronal patterns at 48h APF are revealed by Futsch staining in wild type (A), *fall[1]* (B), and *fall[2]* (C).

D-F'': Overexpression of GFP-Rap1 constitutively active form in IFM. 76h APF hemi-thoraces are stained for GFP and F-actin in control *Actin88f-GAL4/+;fall[1]/+* (D-D''), *Actin88f-GAL4/+; fall[1]* (E-E''), and *Actin88f-GAL4/+;UAS-GFP-Rap1-CA/+; fall[1]* (F-F'').

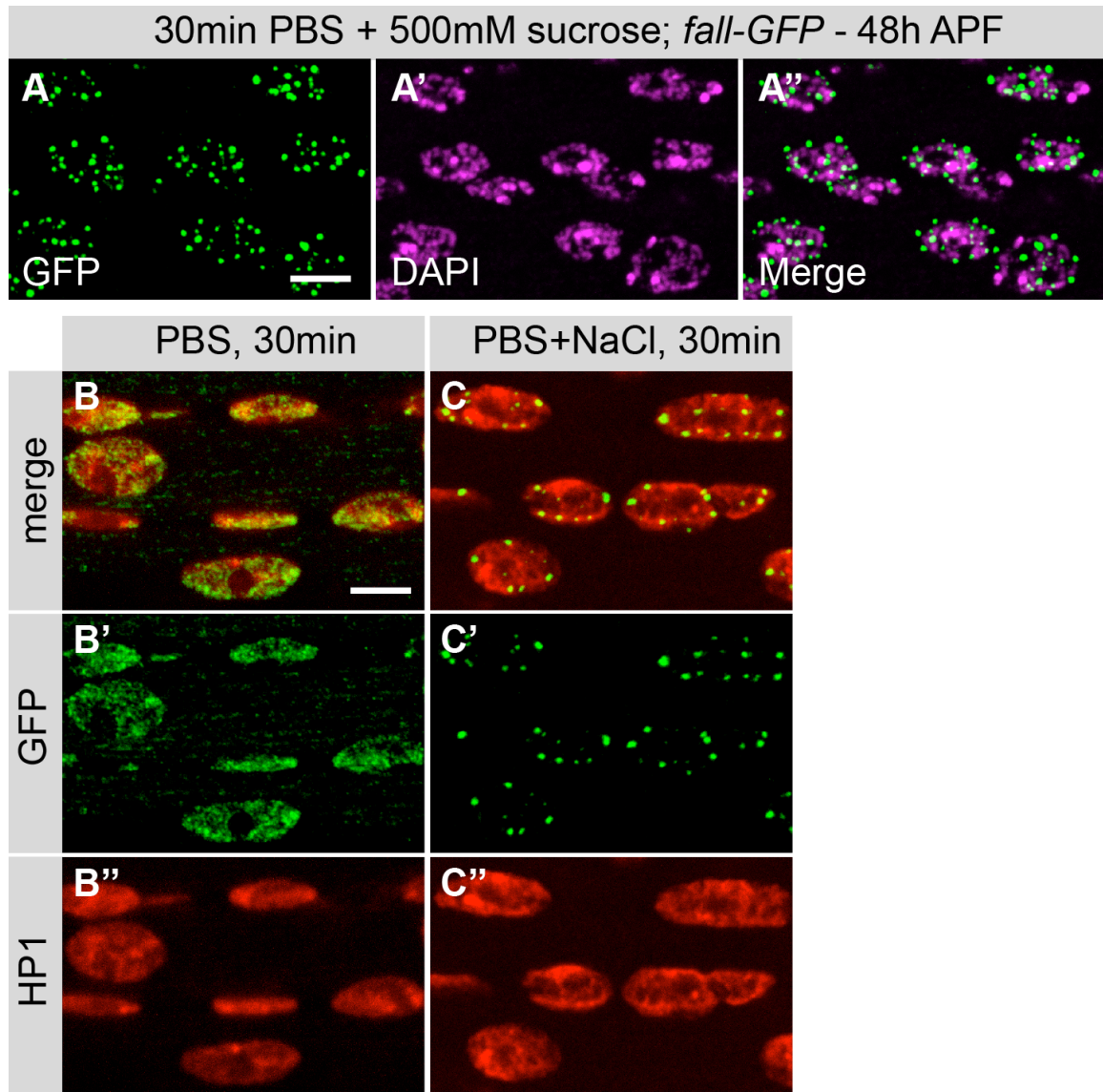
G-H: 76h APF of *Pinch-FL; fall[1]/+* (G) and *Pinch-FL; fall[1]* (H).

Scale bar are 100 μm in all cases. Images of (A-C), (D-F''), and (G-H) are cropped to the same size, respectively.



Supplement Figure 3: Tendon cell debris degraded at 72h APF.

A-D: Dissected 72h APF hemi-thoraces stained for exact the same proteins in Figure 3 (G-N’’’). Tendon debris is absent at muscle tips.

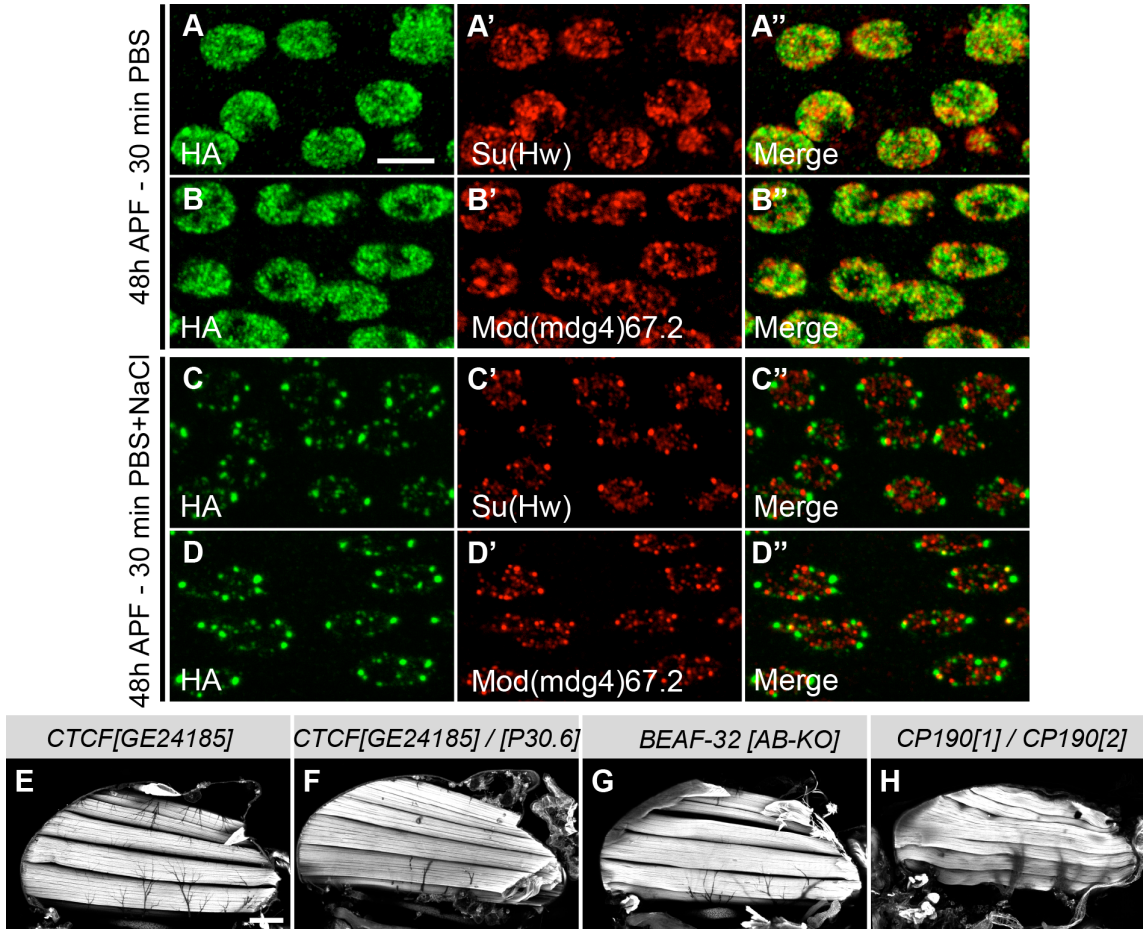


Supplement Figure 4: Fall-GFP response to sucrose treatment; Fall-GFP nuclear body does not co-localize with HP1

A-A'': Fall-GFP localization pattern after sucrose treatment.

B-C'': Fall-GFP nuclear body does not colocalize with HP1.

Scale bars represent 5 μ m in all images. All images are cropped to the same size.



Supplementary Figure 5: Fall-HA does not co-localize with Su(Hw) or Mod(mdg4)67.2; Insulator proteins mutants do not display obvious IFM phenotype.

A-B'': Dissected 48h APF IFMs from *fall-CDS-HA* are stained for HA (A and B) and Su(Hw) (A'), and Mod(mdg4)67.2 (B').

C-D'': Dissected 48h APF IFMs from *fall-CDS-HA* are treated with 250 mM NaCl for 30min and then stained for HA (C and D) and Su(Hw) (C'), and Mod(mdg4)67.2 (D'). In the merged image (C'' and D''), Fall-HA does not colocalize with Su(Hw) or Mod(mdg4)67.2.

E-H: Dissected pharate or 1 day adult hemithoraces of *CTCF[GE24185]* (E), *CTCF[GE24185]CTCF/[P30.6]* (F), *BEAF-32[AB-KO]* (G), and *CP190[1]/CP190[2]* (H). All of them showed intact IFM fibers in the thorax.

Scale bars represent 5 μ m in A-D'' and 100 μ m in E-H. . Images of A-D'' and E-H are cropped to the same size, respectively

References

1. Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *22*, 2008–2017.
2. Bischof, J., Björklund, M., Furger, E., Schertel, C., Taipale, J., and Basler, K. (2013). A versatile platform for creating a comprehensive UAS-ORFeome library in *Drosophila*. *Development* *140*, 2434–2442.
3. Brower, D.L., Wilcox, M., Piovant, M., Smith, R.J., and Reger, L.A. (1984). Related cell-surface antigens expressed with positional specificity in *Drosophila* imaginal discs. *81*, 7485–7489.
4. Brown, N.H. (2000). Cell-cell adhesion via the ECM: integrin genetics in fly and worm. *Matrix Biology : Journal of the International Society for Matrix Biology* *19*, 191–201.
5. Bryantsev, A.L., Baker, P.W., Lovato, T.L., Jaramillo, M.S., and Cripps, R.M. (2012). Differential requirements for Myocyte Enhancer Factor-2 during adult myogenesis in *Drosophila*. *Dev. Biol.* *361*, 191–207.
6. Butcher, R.D.J. (2004). The *Drosophila* centrosome-associated protein CP190 is essential for viability but not for cell division. *J. Cell. Sci.* *117*, 1191–1199.
7. Collier, V.L., Kronert, W.A., O'Donnell, P.T., Edwards, K.A., and Bernstein, S.I. (1990). Alternative myosin hinge regions are utilized in a tissue-specific fashion that correlates with muscle contraction speed. *4*, 885–895.
8. Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* *26*, 1367–1372.
9. Cox, J., Hein, M.Y., Lubner, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* *13*, 2513–2526.
10. Cripps, R.M., Ball, E., Stark, M., Lawn, A., and Sparrow, J.C. (1994). Recovery of dominant, autosomal flightless mutants of *Drosophila melanogaster* and identification of a new gene required for normal muscle structure and function. *Genetics* *137*, 151–164.
11. Cripps, R.M., Suggs, J.A., and Bernstein, S.I. (1999). Assembly of thick filaments and myofibrils occurs in the absence of the myosin head. *The EMBO Journal*.
12. Eberl, H.C., Spruijt, C.G., Kelstrup, C.D., Vermeulen, M., and Mann, M. (2013). A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. *Mol. Cell* *49*, 368–378.
13. Ehler, E., and Gautel, M. (2008). The Sarcomere and Sarcomerogenesis. In *Link.Springer.com*, (New York, NY: Springer New York), pp. 1–14.
14. Ellis, S.J., Goult, B.T., Fairchild, M.J., Harris, N.J., Long, J., Lobo, P., Czerniecki, S., Van Petegem, F., Schöck, F., Peifer, M., et al. (2013). Talin autoinhibition is required for morphogenesis. *Curr. Biol.* *23*, 1825–1833.
15. Fernandes, J.J., Celniker, S.E., and VijayRaghavan, K. (1996). Development of the indirect flight muscle attachment sites in *Drosophila*: role of the PS integrins and the stripe gene. *Dev. Biol.* *176*, 166–184.

Appendices

16. Firdaus, H., Mohan, J., Naz, S., Arathi, P., Ramesh, S.R., and Nongthomba, U. (2015). A cis-regulatory mutation in troponin-I of *Drosophila* reveals the importance of proper stoichiometry of structural proteins during muscle assembly. *Genetics* 200, 149–165.
17. Gómez-Díaz, E., and Corces, V.G. (2014). Architectural proteins: regulators of 3D genome organization in cell fate. *Trends in Cell Biology* 24, 703–711.
18. Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A Phase Separation Model for Transcriptional Control. *Cell* 169, 13–23.
19. Huxley, A.F., and NIEDERGERKE, R. (1954). Structural Changes in Muscle During Contraction: Interference Microscopy of Living Muscle Fibres. *Nature* 173, 971–973.
20. Huxley, H., and Hanson, J. (1954). Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation. *Nature* 173, 973–976.
21. Johnston, J.J., Kelley, R.I., Crawford, T.O., Morton, D.H., Agarwala, R., Koch, T., Schäffer, A.A., Francomano, C.A., and Biesecker, L.G. (2000). A novel nemaline myopathy in the Amish caused by a mutation in troponin T1. *Am. J. Hum. Genet.* 67, 814–821.
22. Kim, J.H., Jin, P., Duan, R., and Chen, E.H. (2015). Mechanisms of myoblast fusion during muscle development. *Curr. Opin. Genet. Dev.* 32, 162–170.
23. Lemke, S.B., and Schnorrer, F. (2016). Mechanical forces during muscle development. *Mech. Dev.*
24. Llewellyn, M., Barretto, R., Delp, S., and Schnitzer, M. (2008). Minimally invasive high-speed imaging of sarcomere contractile dynamics in mice and humans. *Nature* 454, 784–788.
25. Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 153, 320–334.
26. Mitrea, D.M., and Kriwacki, R.W. (2016). Phase separation in biology; functional organization of a higher order. *Cell Communication and Signaling* 2016 14:1 14, 1.
27. Mohan, M., Bartkuhn, M., Herold, M., Philippen, A., Heintl, N., Bardenhagen, I., Leers, J., White, R.A.H., Renkawitz-Pohl, R., Saumweber, H., et al. (2007). The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *The EMBO Journal* 26, 4203–4214.
28. Montana, E.S., and Littleton, J.T. (2004). Characterization of a hypercontraction-induced myopathy in *Drosophila* caused by mutations in Mhc. *J. Cell Biol.* 164, 1045–1054.
29. Moreno-Mateos, M.A., Vejnar, C.E., Beaudoin, J.-D., Fernandez, J.P., Mis, E.K., Khokha, M.K., and Giraldez, A.J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods*.
30. Patel, N., Smith, L.L., Faqeih, E., Mohamed, J., Gupta, V.A., and Alkuraya, F.S. (2014). ZBTB42 mutation defines a novel lethal congenital contracture syndrome (LCCS6). *Hum. Mol. Genet.*
31. Pronovost, S.M., Beckerle, M.C., and Kadrmas, J.L. (2013). Elevated expression of the integrin-associated protein PINCH suppresses the defects of *Drosophila melanogaster* muscle hypercontraction mutants. *PLoS Genet.* 9, e1003406.

32. Reedy, M.C., and Beall, C. (1993). Ultrastructure of Developing Flight Muscle in *Drosophila*. I. Assembly of Myofibrils. *Dev. Biol.* *160*, 443–465.
33. Reedy, M., Bullard, B., and Vigoreaux, J. (2000). Flightin is essential for thick filament assembly and sarcomere stability in *Drosophila* flight muscles. *151*, 1483.
34. Riemer, D., Stuurman, N., Berrios, M., Hunter, C., Fisher, P.A., and Weber, K. (1995). Expression of *Drosophila* lamin C is developmentally regulated: analogies with vertebrate A-type lamins. *J. Cell. Sci.* *108 (Pt 10)*, 3189–3198.
35. Robinson, P., Lipscomb, S., Preston, L.C., Altin, E., Watkins, H., Ashley, C.C., and Redwood, C.S. (2007). Mutations in fast skeletal troponin I, troponin T, and beta-tropomyosin that cause distal arthrogryposis all increase contractile function. *The FASEB Journal* *21*, 896–905.
36. Sanger, J.W., Wang, J., Fan, Y., White, J., and Sanger, J.M. (2010). Assembly and dynamics of myofibrils. *Journal of Biomedicine and Biotechnology* *2010*, 858606.
37. Sarov, M., Barz, C., Jambor, H., Hein, M.Y., Schmied, C., Suchold, D., Stender, B., Janosch, S., KJ, V.V., Krishnan, R.T., et al. (2016). A genome-wide resource for the analysis of protein localisation in *Drosophila*. *Elife* *5*, e12068.
38. Schiaffino, S., and Reggiani, C. (2011). Fiber Types in Mammalian Skeletal Muscles. *Physiol. Rev.* *91*, 1447–1531.
39. Schnorrer, F., and Dickson, B.J. (2004). Muscle building; mechanisms of myotube guidance and attachment site selection. *Dev. Cell* *7*, 9–20.
40. Schnorrer, F., Schönbauer, C., Langer, C.C.H., Dietzl, G., Novatchkova, M., Schernhuber, K., Fellner, M., Azaryan, A., Radolf, M., Stark, A., et al. (2010). Systematic genetic analysis of muscle morphogenesis and function in *Drosophila*. *Nature* *464*, 287–291.
41. Schoborg, T., Rickels, R., Barrios, J., and Labrador, M. (2013). Chromatin insulator bodies are nuclear structures that form in response to osmotic stress and cell death. *J. Cell Biol.* *202*, 261–276.
42. Schönbauer, C., Distler, J., Jährling, N., Radolf, M., Dodt, H.-U., Frasch, M., and Schnorrer, F. (2011). Spalt mediates an evolutionarily conserved switch to fibrillar muscle fate in insects. *Nature* *479*, 406–409.
43. Schweitzer, R., Zelzer, E., and Volk, T. (2010). Connecting muscles to tendons: tendons and musculoskeletal development in flies and vertebrates. *Development* *137*, 2807–2817.
44. Siggs, O.M., and Beutler, B. (2012). The BTB-ZF transcription factors. *Cell Cycle* *11*, 3358–3369.
45. Soler, C., Han, J., and Taylor, M.V. (2012). The conserved transcription factor Mef2 has multiple roles in adult *Drosophila* musculature formation. *Development* *139*, 1270–1275.
46. Sparrow, J.C., and Schöck, F. (2009). The initial steps of myofibril assembly: integrins pave the way. *Nature Reviews Molecular Cell Biology* *10*, 293–298.
47. Spletter, M.L., Barz, C., Yeroslaviz, A., Schonbauer, C., Ferreira, I.R.S., Sarov, M., Gerlach, D., Stark, A., Habermann, B.H., and Schnorrer, F. (2015). The RNA-binding protein Arrest (Bruno) regulates alternative splicing to enable myofibril maturation in *Drosophila* flight muscle. *EMBO Rep.* *16*, 178–191.

Appendices

48. Spletter, M.L., and Schnorrer, F. (2014). Transcriptional regulation and alternative splicing cooperate in muscle fiber-type specification in flies and mammals. *Exp. Cell Res.* *321*, 90–98.
49. Strumpf, D., and Volk, T. (1998). Kakapo, a novel cytoskeletal-associated protein is essential for the restricted localization of the neuregulin-like factor, vein, at the muscle-tendon junction site. *143*, 1259–1270.
50. Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *6*, e21800.
51. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* *13*, 731–740.
52. Vakaloglou, K.M., Chrysanthis, G., and Zervas, C.G. (2016). IPP Complex Reinforces Adhesion by Relaying Tension-Dependent Signals to Inhibit Integrin Turnover. *Cell Reports*.
53. Venken, K.J.T., Schulze, K.L., Haelterman, N.A., Pan, H., He, Y., Evans-Holm, M., Carlson, J.W., Levis, R.W., Spradling, A.C., Hoskins, R.A., et al. (2011). MiMIC: a highly versatile transposon insertion resource for engineering *Drosophila melanogaster* genes. *Nat. Methods* *8*, 737–743.
54. Vigoreaux, J.O. (2006). Molecular Basis of Muscle Structure. In *Muscle Development in Drosophila*, (New York, NY: Springer New York), pp. 143–156.
55. Weitkunat, M., and Schnorrer, F. (2014). A guide to study *Drosophila* muscle biology. *68*, 2–14.
56. Weitkunat, M., Kaya-Çopur, A., Grill, S.W., and Schnorrer, F. (2014). Tension and force-resistant attachment are essential for myofibrillogenesis in *Drosophila* flight muscle. *Curr. Biol.* *24*, 705–716.
57. Zambon, A.C., Gaj, S., Ho, I., Hanspers, K., Vranizan, K., Evelo, C.T., Conklin, B.R., Pico, A.R., and Salomonis, N. (2012). GO-Elite: a flexible solution for pathway and ontology over-representation. *28*, 2209–2210.
58. Zhang, X., Ferreira, I.R.S., and Schnorrer, F. (2014a). A simple TALEN-based protocol for efficient genome-editing in *Drosophila*. *Methods* *69*, 32–37.
59. Zhang, X., Koolhaas, W.H., and Schnorrer, F. (2014b). A versatile two-step CRISPR-and RMCE-based strategy for efficient genome engineering in *Drosophila*. *G3 (Bethesda)* *4*, 2409–2418.
60. Zhu, L., and Brangwynne, C.P. (2015). Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Current Opinion in Cell Biology* *34*, 23–30.

Acknowledgement

First of all, I would like to express my sincere and deepest gratitude to my supervisor Dr. Frank Schnorrer for his constant scientific support during my Ph.D. study as well as his personal support to my family during our stay. Without his encouragement during all stages of the thesis, including support for me to present my projects at scientific meetings and participate in prestigious training courses, and most importantly his insightful and sharp inspiration to my project, I would not be able to accomplish my Ph.D.

I am sincerely thankful to Prof. Dr. Barbara Conradt for being my first referee of my thesis, Prof. Nicolas Gompel, Prof. Dr. John Parsch, Prof. Angelika Boettger, and my supervisor Dr. Frank Schnorrer for being my thesis examiners as well as their time and dedication to review my thesis.

I am sincerely grateful to Dr. Bianca Habermann for her help with the bioinformatics analysis of the *fall* project, which resulted in a successful collaboration.

I am also thankful to my TAC committee members Dr. Juerg Mueller, Dr. Anne K. Classen, and Dr. Carla E. Margulies for their useful input and valuable advices to my project.

I was lucky to be accepted to the IMPRS-LS program and spent impressive years with my peers. I would like to express my thank to the IMPRS coordinators Dr. Hans Joerg Schaeffer, Dr. Ingrid Wolf, Ms. Maximiliane Reif, and Dr. Amy Gerc for taking care of me in many aspects including numerous great workshops, retreats, my VISA application, my TAC meeting, university registration...

I especially want to thank the former and present members of the Schnorrer lab for their scientific input during lab meetings, for the enjoyable and impressive lab life, and for their help.

I also want to express my appreciation to Prof. Dr. Klaus Foerstemann for sharing CRISPR reagents, Dr. Phillip Port for sharing CRISPR flies, Prof. Dr. Talila Volk and Prof. Dr. Mariano Labrador for antibodies.

Lastly, I am deeply grateful to my parents, and my brother for their support, patience, and encouragement during the past years. In particular, my deepest gratitude and love goes to my wife Shasha for her love, constant support, patience, and for taking care our baby Muchen.

Curriculum Vitae

Curriculum Vitae

Personal info

Name: Xu Zhang Gender: Male

Date of Birth: 28. Nov. 1986

Education

2012-present	IMPRS-LS Ph.D. Candidate Max Planck Institute of Biochemistry
2009-2012	Master, South China University of Technology
2005-2009	Bachelor, Harbin Institute of Technology

Awards and Honors

2009	Qualified to the Master Program without Examination
2009	Outstanding Graduate of Hei Long Jiang Province (1 in our College)
2009	Outstanding Graduate of Harbin Institute of Technology (C9 League University)
2009	Excellent Bachelor Thesis of Harbin Institute of Technology (1 in Our College)
2009	National Top Grade Scholarship
2010	National Top Grade Scholarship
2012	Excellent Master Thesis of South China University of Technology

Publications

1. **Zhang, X.**, Schnorrer, F. fall, a BTB-Zinc finger protein, is essential for sarcomere genesis to prevent hypercontraction in skeletal muscle. Manuscript ready for submission .
2. **Zhang, X.**, Schnorrer, F. (2017) AIDing targeted protein degradation in *Drosophila*. *The FEBS Journal*, Accepted.
3. **Zhang, X.**, Koolhaas, W.H., and Schnorrer, F. (2014) A versatile two-step CRISPR- and RMCE-based strategy for efficient genome engineering in *Drosophila*. *G3*, 4(12), 2409–2418.
4. **Zhang, X.**, Ferreira, Irene R.S., and Schnorrer, F. (2014). A simple TALEN-based protocol for efficient genome-editing in *Drosophila*. *Methods* 69, 32-37.
5. **Zhang, X.**, Chen, G., Qin, CY., Wang, YL., and Wei, D. (2012). Slr0643, an S2P homologue, is essential for acid acclimation in the cyanobacterium *Synechocystis* sp. PCC 6803. *Microbiology-SGM* 158, 2765-2780.
6. Chen, G., and **Zhang, X.** (2010). New insights into S2P signaling cascades: regulation, variation, and conservation. *Protein Sci* 19, 2015-2030.

7. Chen, G., Wang, H., **Zhang, X.**, and Yang, S.T. (2014). Nutraceuticals and functional foods in the management of hyperlipidemia. *Crit Rev Food Sci*, 54(9), 1180–1201.
8. Chen, G., Law, K., Ho, P., **Zhang, X.**, and Li, N. (2012). EGY2, a chloroplast membrane metalloprotease, plays a role in hypocotyl elongation in *Arabidopsis*. *Mol Biol Rep* 39, 2147-2155.

Conferences

- **23rd European Drosophila Research Conference, Barcelona, Spain, 2013**
Oral Presentation: TALEN Mediated Gene Targeting in *Drosophila*.
- **Regional Drosophila Meeting, Heidelberg, Germany, 2014**
Poster Presentation: A Versatile CRISPR-Cas9 and RMCE Based Genome-engineering Platform in *Drosophila*.
- **Drosophila Genetics and Genomics Course, Cambridge, United Kingdom, 2014**
Oral Presentation: A Practical Guideline for the Application of CRISPR-Cas9 Based Genome-engineering Platform in *Drosophila*.
- **Joint Meeting of the German and French Societies of Developmental Biologists, Nuremberg, Germany, 2015**
Poster Presentation: A Practical Guideline for Using CRISPR-Cas9 and RMCE Based Genome-Engineering Platform in *Drosophila*.
- **Gordon Research Conference --- Myogenesis, Lucca, Italy, 2015**
Poster Presentation: *fallen angel* encodes a BTB/POZ zinc-finger protein essential for *Drosophila* flight muscle integrity.

References

Prof. Dr. Gu Chen

Email: guchen@scut.edu.cn

School of Food Science and Engineering

South China University of Technology

Dr. Frank Schnorrer

Email: frank.SCHNORRER@univ-amu.fr

Muscle Dynamics Group

IBDM – Institut de Biologie du Développement de Marseille (UMR 7288)

CNRS & Aix-Marseille Université